



Pós-graduação em Ciência da Computação

Bruno Iran Ferreira Maciel

Otimização de Falsos Alarmes em Detecção de Mudança de Conceito



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
<http://cin.ufpe.br/~posgraduacao>

Recife

2020

Bruno Iran Ferreira Maciel

Otimização de Falsos Alarmes em Detecção de Mudança de Conceito

Tese de Doutorado apresentada ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Ciência da Computação.

Área de Concentração: Inteligência Computacional

Orientador: Prof. Roberto Souto Maior de Barros

Recife

2020

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

M152o Maciel, Bruno Iran Ferreira
Otimização de Falsos Alarmes em Detecção de Mudança de Conceito/ Bruno
Iran Ferreira Maciel. – 2020.
260 f.: il., fig., tab.

Orientador: Roberto Souto Maior de Barros
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da
Computação, Recife, 2020.
Inclui referências e apêndices.

1. Inteligência computacional. 2. Fluxo de dados. I. Barros, Roberto Souto
Maior de (orientador). II. Título.

006.31

CDD (23. ed.)

UFPE - CCEN 2020 - 136

Bruno Iran Ferreira Maciel

“Otimização de Falsos Alarmes em Detecção de Mudança de Conceito”

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação.

Aprovado em: 04/03/2020.

Prof. Roberto Souto Maior de Barros

BANCA EXAMINADORA

Prof. Dr. Paulo Maurício Gonçalves Júnior
Instituto Federal de Pernambuco / Campus Recife

Prof. Dr. Germano Crispim Vasconcelos
Centro de Informática / UFPE

Prof. Dr. Wilson Rosa de Oliveira Júnior
Departamento de Estatística e Informática / UFRPE

Prof. Dr. Jean Paul Barddal
Programa de Pós-Graduação em Informática / PUC/PR

Prof. Dr. Francisco Madeiro Bernardino Júnior
Centro de Ciências e Tecnologia / UNICAP

Dedico esta tese às pessoas mais importantes para mim, minha querida mãe Márcia da Costa Maciel, meu pai Ronaldo Ferreira Maciel, ela e ele são responsáveis por minha vida e a quem devo meu caráter e disciplina. Além dela e dele, também à adorada companheira e amada, Sylvia Emmanuelle Castelo Branco de Holanda Victor, sempre paciente e generosa em meus momentos de desânimos e falta de estímulos. Ao mesmo tempo, estendo esta dedicatória a minha irmã à Dra. Jackeline da Costa Maciel. Ofereço este trabalho em memória de minha segunda mãe Maria de Fátima Castelo Branco de Holanda Victor.

AGRADECIMENTOS

Agradeço primeiramente a Deus que está sempre ao nosso lado. Meus pais Márcia e Ronaldo, minhas irmãs Jackeline, Roberta, meu irmão Fábio e meus amigos do Janga, pessoas que me apoiam e me incentivam a realizar meus sonhos. Sempre serei grato a todos vocês. Principalmente a minha mãe que tanto tem me apoiado em todos os momentos e minha irmã Jackeline juntamente com seu esposo Humberto Silva que sempre mostram-se solidários e dispostos a me ajudar. A todos os meus amigos pelo companheirismo e pelas horas de descontração (raríssimas fora do CIn-UFPE). Sempre me lembrarei de todos vocês. Obrigado a todos.

Ao professor Dr. Roberto Souto Maior de Barros, por todo seu caráter, orientações neste trabalho, profissionalismo, amizade, paciência, me ajudar a corrigir falhas e dedicação comigo durante as diversas atividades de pesquisa e andamento deste trabalho.

Ao amigo Dr. Silas Garrido Teixeira de Carvalho Santos e minha querida Sylvia Victor por contribuírem para meu amadurecimento como pesquisador e principalmente pela disponibilidade e solidariedade em me auxiliar nas revisões e escrita do texto final da tese. Além disso, estendo os agradecimentos aos amigos Juan Isidro Gonzalez, Jose Luis Martinez Perez, Rohgi Toshio Meneses Chikuhi e Danilo Rafael de Lima Cabral pelas parcerias e apoio nas questões teóricas, escrita de artigos e vida pessoal.

Agradecimento especial aos amigos Webber de Souza Fantini e Danilo Monteiro Ribeiro. Ambos contribuíram para a continuidade da tese ao me indicarem oportunidades de trabalho como professor, desta forma foi possível financiar minhas despesas mantendo dedicação, tempo e atenção exigida.

Aos amigos que fiz no Centro de Informática/UFPE durante os anos de preparação, a saber: Danusa Ribeiro Bezerra da Cunha, Aline Chagas Rodrigues Marques, Sidartha Azevedo Lobo de Carvalho, Paulo Henrique, Webber Fantini, Márcio Angelo, René, Hilário Tomás, Andriano, Brunno Wagner, Halisson Cardoso, Ruan, Caio Barbosa, Elaine de Assis, Carla Cristina Braz de Oliveira, Myller Claudino de Freitas e Emanuel Carlos Gomes Ferraz Silva. Todavia, também dedico aos demais amigos/colegas que não mencionei os nomes aqui.

Por fim, agradeço ao CIn-UFPE pela oportunidade e CAPES pelo apoio financeiro que viabilizou o trabalho. Também sou muito grato a todos os colaboradores do CIn-UFPE.

RESUMO

Mudança de Conceito é um problema inerente ao aprendizado em fluxo de dados. A detecção de mudanças de conceito é uma técnica conhecida por rastrear o ponto de mudança entre conceitos, e os métodos que implementam essa ideia geralmente são conhecidos como detectores. Esta tese é dedicada ao estudo de algoritmos de detecção de mudanças de conceito. Inicialmente foi feita uma revisão da literatura, incluindo os principais trabalhos relacionados. Na sequência, foram feitas contribuições ao *Massive Online Analysis* (MOA) com complementos de códigos, incluindo a ferramenta *Massive Online Analysis Manager PHP framework* (MOAManager), desenvolvida e utilizada para gerenciar experimentos usando o MOA. Em seguida, foram comparados os desempenhos dos trabalhos relacionados, separando-os em detectores de única e de múltipla análise. Posteriormente, são apresentados os métodos propostos de comitês de detectores *Drift Detection Ensemble* (DDE), *Mixed Signals Drift Detection Ensemble* (MSDDE) e *Match Drift Detection Ensemble* (MDDE), assim como os métodos convencionais de detecção *Ultimately Simple Drift Detector* (USDD) e *USDD eXtension* (USDDX). Nesta tese foram também realizadas comparações de larga escala com 25 detectores, 6 geradores de dados, usando três tamanhos diferentes de bases de dados com simulações de mudanças abruptas e graduais, além de 7 bases de dados reais, todos utilizando dois classificadores (*Naive Bayes* e *Hoeffding Tree*). Os resultados empíricos mostram que os comitês de detectores propostos foram competitivos em relação aos trabalhos do estado da arte em detecção de mudança de conceito, oferecendo resultados melhores em vários casos em comparação aos trabalhos de única análise e sendo superior na maioria das vezes em relação aos métodos de múltipla análise. Ambos os métodos, USDD e USDDX, foram competitivos em relação aos trabalhos relacionados, apresentando-se como alternativas robustas para detecção de mudanças de conceito. Finalmente, na maioria dos cenários avaliados, os métodos propostos ofereceram bons resultados em relação à acurácia *Prequencial* e principalmente na redução de falsos alarmes de detecção em transições abruptas, mostrando-se como uma alternativa competitiva.

Palavras-chaves: Mudança de conceito. Comitê. Fluxo de dados. Aprendizado *on-line*.

ABSTRACT

Concept drift is an inherent problem of learning in data stream environments. Concept Drift Detection is a technique known to track the points of change between concepts, and methods that implement this idea are often referred to detectors. This thesis is dedicated to the study of concept drift detection algorithms. Initially, a literature review was performed, surveying the main related works. Subsequently, contributions were made to the MOA framework with code implementations, including the MOAManager tool, developed and used to manage experiments run using MOA. In addition, the performances of the related works were compared, separating them into single and multiple analysis detectors. Then, the proposed methods of detector ensembles DDE, MSDDE and MDDE, as well as the conventional detection methods USDD and USDDX, are presented. In this thesis, large-scale comparisons were also performed with 25 detectors, 6 dataset generators, using three different databases sizes with abrupt and gradual changes simulations, as well as seven real-world datasets, all using two base classifiers (*Naive Bayes* and *Hoeffding Tree*). Empirical results shows that the proposed detector ensembles were competitive with state-of-the-art works on concept drift detection, providing better results in several cases when compared to single-analysis works, and being superior to multiple analysis methods in most cases. Both USDD and USDDX were competitive in relation to related works, presenting themselves as robust alternatives for detecting changes of concept. Finally, in most of the evaluated scenarios, the proposed methods offered good results related to the Prequential accuracy and, especially, in the reduction of false detection alarms in abrupt transitions, proving to be competitive alternatives.

Keywords: Concept Drift. Ensemble. Data stream. Online learning.

LISTA DE FIGURAS

| | |
|---|-----|
| Figura 1 – Exemplo de mudança de conceito. Seja S_1 e S_2 conceitos, T é dado pela mistura entre eles e X_1, \dots, X_t as microinformações do domínio do conceito. | 43 |
| Figura 2 – Exemplo de mudança de conceito implementada no framework MOA. Seja S_1 e S_2 conceitos, T é dado pelo período de transição entre conceitos. A mistura entre conceitos, expresso por T , corresponde ao período no qual há dois conceitos S_1 e S_2 , ou seja, duas distribuições de variáveis. | 44 |
| Figura 3 – Agente. Abstração de sistema básico de entrada e saída baseado em atuador e sensor. O objetivo é fazer um comparativo entre os papéis do sistema e algoritmo de adaptação, sensor e algoritmo de Detecção de Mudança de Conceito (DMC), assim como o atuador em relação ao algoritmo de Aprendizado de Máquina (do inglês, <i>Machine Learning</i>) (AM). As entradas x podem ser comparadas aos dados em Aprendizado Adaptativo com Detector Embutido (AADE) e saídas y como os resultados das classificações. | 58 |
| Figura 4 – Abordagem de Aprendizado Adaptativa com Detector Embutido. . . . | 59 |
| Figura 5 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise. . . . | 84 |
| Figura 6 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise. . . . | 85 |
| Figura 7 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise. . . . | 86 |
| Figura 8 – Resumo das diferenças calculadas com acurácia, <i>Mean Time Ratio</i> (traduzindo para o português, Taxa Média de Tempo) (MTR) e <i>Matthews Correlation Coefficient</i> (traduzindo para o português, Coeficiente de Correlação de Matthews) (MCC) para HT e NB. Considerando comparações com os métodos propostos de múltipla análise. | 88 |
| Figura 9 – Resumo das vitórias calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com os métodos propostos de múltipla análise. | 89 |
| Figura 10 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. . . . | 106 |
| Figura 11 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise. . . . | 110 |

| | |
|---|-----|
| Figura 12 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise. . . . | 113 |
| Figura 13 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com os métodos propostos de única análise. | 115 |
| Figura 14 – Correlação de eficiência do <i>Fast Hoeffding Drift Detection Method</i> (FHDDM) com HT e NB. | 117 |
| Figura 15 – Correlação de eficiência do <i>Stacking Fast Hoeffding Drift Detection Method</i> (FHDDMS) com HT e NB. | 117 |
| Figura 16 – Correlação de eficiência do <i>Fisher Proportions Drift Detector</i> (FPDD) com HT e NB. | 118 |
| Figura 17 – Correlação de eficiência do <i>Fisher Square Drift Detector</i> (FSDD) com HT e NB. | 118 |
| Figura 18 – Correlação de eficiência do <i>Fisher Test Drift Detector</i> (FTDD) com HT e NB. | 118 |
| Figura 19 – Correlação de eficiência do <i>Hoeffding-based Drift Detection Method A-Test</i> (HDDMA) com HT e NB. | 119 |
| Figura 20 – Correlação de eficiência do <i>Hoeffding-based Drift Detection Method W-Test</i> (HDDMW) com HT e NB. | 119 |
| Figura 21 – Correlação de eficiência do <i>Reactive Drift Detection Method</i> (RDDM) com HT e NB. | 119 |
| Figura 22 – Correlação de eficiência do <i>Wilcoxon Rank Sum Test Drift Detector</i> (WSTD) com HT e NB. | 120 |
| Figura 23 – Correlação de eficiência do USDD com HT e NB. | 120 |
| Figura 24 – Correlação de eficiência do USDDX com HT e NB. | 120 |
| Figura 25 – MOAManager Workflow | 124 |
| Figura 26 – Ilustração de como a MOAManager pode ser usada para gerar <i>Scripts</i> MOA. | 125 |
| Figura 27 – Ilustra como os dados podem ser extraídos. | 128 |
| Figura 28 – Exemplos de diagramas gerados por MOAManager referentes a resultados de testes estatísticos. | 129 |
| Figura 29 – MOAManager integração com MOA | 130 |
| Figura 30 – Zona de tolerância na detecção | 156 |
| Figura 31 – Error tipo 1 e erro tipo 2 | 158 |
| Figura 32 – Escala do Coeficiente de Correlação de Matthews | 159 |
| Figura 33 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados. | 167 |

| | |
|---|-----|
| Figura 34 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados. | 170 |
| Figura 35 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados. | 173 |
| Figura 36 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando as comparações com os métodos de única análise dos trabalhos relacionados. | 174 |
| Figura 37 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 182 |
| Figura 38 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 183 |
| Figura 39 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 184 |
| Figura 40 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando as comparações com os métodos de múltipla análise dos trabalhos relacionados. | 186 |
| Figura 41 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do DDE. | 201 |
| Figura 42 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do DDE. | 202 |
| Figura 43 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do DDE. | 204 |
| Figura 44 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do DDE. | 205 |
| Figura 45 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do MSDDE. | 206 |
| Figura 46 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do MSDDE. | 207 |
| Figura 47 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do MSDDE. | 209 |
| Figura 48 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do MSDDE. | 210 |
| Figura 49 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do MDDE. | 211 |

| | |
|--|-----|
| Figura 50 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do MDDE. | 212 |
| Figura 51 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do MDDE. | 214 |
| Figura 52 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do MDDE. . . | 215 |

LISTA DE QUADROS

| | |
|---|-----|
| Quadro 1 – Categorias dos estados de detecção de mudança de conceito. | 58 |
| Quadro 2 – Resumo das camadas do Modelo de referência de AADE. | 58 |
| Quadro 3 – Matriz de confusão com resultados de detecção | 154 |

LISTA DE TABELAS

| | |
|--|-----|
| Tabela 1 – Resumo descritivo dos geradores. No gerador RRBF a quantidade de conceitos é representada por n que pode variar de 1 à maior que 1 milhão. | 54 |
| Tabela 2 – Resumo descritivo das bases de dados reais. | 55 |
| Tabela 3 – Resumo técnico das estações de trabalho utilizadas para realização dos experimentos. | 66 |
| Tabela 4 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise. | 90 |
| Tabela 5 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise. | 91 |
| Tabela 6 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise. | 92 |
| Tabela 7 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise. | 93 |
| Tabela 8 – Médias dos MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise. | 94 |
| Tabela 9 – Médias dos MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise. | 95 |
| Tabela 10 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 104 |
| Tabela 11 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 105 |
| Tabela 12 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 108 |
| Tabela 13 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 109 |
| Tabela 14 – Médias dos MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 111 |
| Tabela 15 – Médias dos MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 112 |
| Tabela 16 – Base de dados com 10k instâncias e 4 drifts reais - situação ideal | 154 |
| Tabela 17 – Base de dados com 10k instâncias e 4 drifts reais | 155 |

| | |
|--|-----|
| Tabela 18 – Base de dados com 10k instâncias e 4 drifts reais - dados estimados e reais de mudança de conceito. A classificação corresponde a categoria na qual a identificação do drift foi realizada, pertencendo a um conjunto de 3 classes, a saber: <i>True Positive</i> (traduzindo para o português, Verdadeiro Positivo) (TP), <i>False Positive</i> (traduzindo para o português, Falso Positivo) (FP) e <i>False Negative</i> (traduzindo para o português, Falso Negativo) (FN). A classificação para <i>True Negative</i> (traduzindo para o português, Verdadeiro Negativo) (TN) corresponde ao total de instâncias na base de dados menos o total de TP + FP + FN. | 156 |
| Tabela 19 – Matriz de confusão com dados da Tabela 18 | 156 |
| Tabela 20 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados. | 165 |
| Tabela 21 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados. | 166 |
| Tabela 22 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados. | 168 |
| Tabela 23 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados. | 169 |
| Tabela 24 – Médias calculadas do MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados. | 171 |
| Tabela 25 – Médias calculadas do MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados. | 172 |
| Tabela 26 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados. | 175 |
| Tabela 27 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados. | 175 |
| Tabela 28 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados. | 176 |

| | |
|---|-----|
| Tabela 29 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados. | 176 |
| Tabela 30 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados. | 177 |
| Tabela 31 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados. | 177 |
| Tabela 32 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 187 |
| Tabela 33 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 188 |
| Tabela 34 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 189 |
| Tabela 35 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 190 |
| Tabela 36 – Médias calculadas do MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 191 |
| Tabela 37 – Médias calculadas do MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 192 |
| Tabela 38 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 193 |
| Tabela 39 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 194 |
| Tabela 40 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 195 |
| Tabela 41 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 196 |

| | |
|--|-----|
| Tabela 42 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 197 |
| Tabela 43 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados. | 198 |
| Tabela 44 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do DDE. | 216 |
| Tabela 45 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do DDE. | 217 |
| Tabela 46 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do DDE. | 218 |
| Tabela 47 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do DDE. | 219 |
| Tabela 48 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do DDE. | 220 |
| Tabela 49 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do DDE. | 221 |
| Tabela 50 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do MSDDE. | 222 |
| Tabela 51 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do MSDDE. | 223 |
| Tabela 52 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do MSDDE. | 224 |
| Tabela 53 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do MSDDE. | 225 |
| Tabela 54 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do MSDDE. | 226 |
| Tabela 55 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do MSDDE. | 227 |
| Tabela 56 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do MDDE. | 228 |
| Tabela 57 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do MDDE. | 229 |
| Tabela 58 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do MDDE. | 230 |
| Tabela 59 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do MDDE. | 231 |

| | |
|---|-----|
| Tabela 60 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do MDDE. | 232 |
| Tabela 61 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do MDDE. | 233 |
| Tabela 62 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do DDE. | 234 |
| Tabela 63 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do DDE. | 234 |
| Tabela 64 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do DDE. | 235 |
| Tabela 65 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do DDE. | 235 |
| Tabela 66 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do DDE. | 236 |
| Tabela 67 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do DDE. | 236 |
| Tabela 68 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do MSDDE. | 237 |
| Tabela 69 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do MSDDE. | 237 |
| Tabela 70 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do MSDDE. | 238 |
| Tabela 71 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do MSDDE. | 238 |
| Tabela 72 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do MSDDE. | 239 |
| Tabela 73 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do MSDDE. | 239 |
| Tabela 74 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do MDDE. | 240 |
| Tabela 75 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do MDDE. | 240 |
| Tabela 76 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do MDDE. | 241 |
| Tabela 77 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do MDDE. | 241 |
| Tabela 78 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do MDDE. | 242 |

| | |
|--|-----|
| Tabela 79 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do MDDE. | 242 |
| Tabela 80 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise. | 243 |
| Tabela 81 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise. | 244 |
| Tabela 82 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise. | 245 |
| Tabela 83 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise. | 246 |
| Tabela 84 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise. | 247 |
| Tabela 85 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise. | 248 |
| Tabela 86 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 249 |
| Tabela 87 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 250 |
| Tabela 88 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 251 |
| Tabela 89 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 252 |
| Tabela 90 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 253 |
| Tabela 91 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 254 |

| | |
|---|-----|
| Tabela 92 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 255 |
| Tabela 93 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise. | 256 |
| Tabela 94 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 257 |
| Tabela 95 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise. | 258 |
| Tabela 96 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 259 |
| Tabela 97 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise. | 260 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------------------|--|
| <i>e</i> -Detector | <i>Selective Detector Ensemble.</i> |
| AADE | Aprendizado Adaptativo com Detector Embutido. |
| AD | <i>All Detectors Detect Drift.</i> |
| ADWIN | <i>ADaptive WINdowing.</i> |
| ALHD | <i>At Least Half of the Detectors Detect Drift.</i> |
| ALO | <i>At Least One Detects Drift.</i> |
| AM | Aprendizado de Máquina (do inglês, <i>Machine Learning</i>). |
| ARF | <i>Adaptive Random Forest.</i> |
| ARL(θ) | <i>Average Run Length</i> (traduzindo para o português, Média de Duração da Execução). |
| BOLE | <i>Boosting-like Online Learning Ensemble.</i> |
| BW | <i>Basic Window.</i> |
| CSV | <i>Comma-Separated Values.</i> |
| CVFDT | <i>Concept-adapting Very Fast Decision Tree learner.</i> |
| DDE | <i>Drift Detection Ensemble.</i> |
| DDM | <i>Drift Detection Method.</i> |
| DMC | Detecção de Mudança de Conceito. |
| DWM | <i>Dynamic Weighted Majority.</i> |
| ECDD | <i>EWMA for Concept Drift Detection.</i> |
| EDDM | <i>Early Drift Detection Method.</i> |
| EWMA | <i>Exponentially Weighted Moving Average.</i> |
| F1 score | F-Measure. |

| | |
|--------|--|
| FF | <i>Fading Factors.</i> |
| FHDDM | <i>Fast Hoeffding Drift Detection Method.</i> |
| FHDDMS | <i>Stacking Fast Hoeffding Drift Detection Method.</i> |
| FN | <i>False Negative</i> (traduzindo para o português, Falso Negativo). |
| FP | <i>False Positive</i> (traduzindo para o português, Falso Positivo). |
| FPDD | <i>Fisher Proportions Drift Detector.</i> |
| FSDD | <i>Fisher Square Drift Detector.</i> |
| FTDD | <i>Fisher Test Drift Detector.</i> |
| GUI | <i>Graphical User Interface</i> (traduzindo para o português, Interface Gráfica de Usuário). |
| HAT | <i>Hoeffding Adaptive Tree.</i> |
| HATT | <i>Hoeffding Anytime Tree.</i> |
| HDDM | <i>Hoeffding-based Drift Detection Method.</i> |
| HDDMA | <i>Hoeffding-based Drift Detection Method A-Test.</i> |
| HDDMW | <i>Hoeffding-based Drift Detection Method W-Test.</i> |
| HT | <i>Hoeffding Tree.</i> |
| HTML | <i>HyperText Markup Language.</i> |
| IID | Independent and Identically Distributed (traduzindo para o português, Independente e Identicamente Distribuída). |
| JAR | <i>JavaTM Archive.</i> |
| JVM | <i>Java Virtual Machine.</i> |
| KUE | <i>Kappa Updated Ensemble.</i> |
| LED | <i>Light Emitting Diode.</i> |
| MC | Mudança de Conceito (do inglês, <i>Concept Drift</i>). |

| | |
|------------|---|
| MCC | <i>Matthews Correlation Coefficient</i> (traduzindo para o português, Coeficiente de Correlação de Matthews). |
| MD | Mineração de Dados (do inglês, <i>Data Mining</i>). |
| MDDE | <i>Match Drift Detection Ensemble</i> . |
| MDR | <i>Missed Detection Rate</i> (traduzindo para o português, Taxa de Detecção Perdida). |
| MOA | <i>Massive Online Analysis</i> . |
| MOAManager | <i>Massive Online Analysis Manager PHP framework</i> . |
| MSDDE | <i>Mixed Signals Drift Detection Ensemble</i> . |
| MTD | <i>Mean Time to Detection</i> (traduzindo para o português, Tempo Médio para Detecção). |
| MTFA | <i>Mean Time between False Alarms</i> (traduzindo para o português, Tempo Médio Entre Falsos Alarmes). |
| MTR | <i>Mean Time Ratio</i> (traduzindo para o português, Taxa Média de Tempo). |
| NB | <i>Naive Bayes</i> . |
| OABM1 | <i>Online AdaBoost-based M1</i> . |
| PAC | <i>Probably Approximately Correct</i> . |
| PHP | <i>PHP: Hypertext Preprocessor</i> . |
| RDDM | <i>Reactive Drift Detection Method</i> . |
| RRBF | <i>Random Radial Basic Function</i> . |
| SEED | <i>Detecting Volatility Shift in Data Streams</i> . |
| SeqDrift1 | <i>Sequential Drift Change Detector 1</i> . |
| SeqDrift2 | <i>Sequential Drift Change Detector 2</i> . |
| SGBD | Sistema de Gerenciamento de Banco de Dados. |
| STEPD | <i>Statistical Test of Equal Proportions</i> . |

| | |
|-------|--|
| SW | <i>Sliding Window.</i> |
| TN | <i>True Negative</i> (traduzindo para o português, Verdadeiro Negativo). |
| TP | <i>True Positive</i> (traduzindo para o português, Verdadeiro Positivo). |
| USDD | <i>Ultimately Simple Drift Detector.</i> |
| USDDX | USDD <i>eXtension.</i> |
| WSTD | <i>Wilcoxon Rank Sum Test Drift Detector.</i> |

NOTAÇÕES E LISTA DE SÍMBOLOS

| | |
|-----------------------|--|
| * | Em geral, variáveis aleatórias são representadas por letras maiúsculas e valores possíveis das variáveis aleatórias são representadas por letras minúsculas. |
| * | Letras gregas são usadas para denotar parâmetros. |
| * | Designa-se por $y_{s:t}$ o conjunto de informação disponível no período s a t , ou seja, $y_{s:t} = (y_s, y_{s+1}, \dots, y_{t-1}, y_t)$. |
| * | Será usada de forma indiferenciada o termo densidade para designar uma função de probabilidade ou uma função densidade de probabilidade. |
| $Pr(\cdot)$ | Usado para representar uma função de probabilidade ou uma função densidade de probabilidade, consoante o contexto; $p(x, y)$ representa a função de probabilidade (densidade de probabilidade) conjunta dos vetores aleatórios \mathbf{X} e \mathbf{Y} ; $p(y x)$ representa função de probabilidade (densidade de probabilidade) condicionada de \mathbf{Y} , dado o vetor \mathbf{X} . |
| $\Pi(\theta)$ | Distribuição a priori de θ . |
| $\Pi(\theta x)$ | Distribuição a posteriori de θ dado x . |
| $f : x \rightarrow y$ | Função f , no qual x e y representam conjuntos de entradas e saída, respectivamente. |
| ρ | Valor de probabilidade |
| $Pr(x, y)$ | Função de distribuição de probabilidade conjunta. |
| $Pr(x \vee y)$ | Forma normal disjuntiva. Regra da soma. |
| $Pr(x \wedge y)$ | Forma normal conjuntiva. Regra do produto. |
| $Pr(a b)$ | A probabilidade condicional de a dado b . |
| $Pr(a b)$ | $Pr(a = a b = b)$. |
| \in | Pertence |
| δ | Delta |
| θ | Teta |

| | |
|----------------|-----------------------------|
| β | Poder do teste. |
| α | Nível de Significância. |
| H_0 | Hipótese nula. |
| H_1 | Hipótese alternativa. |
| H_2 | Hipótese alternativa. |
| \hat{C} | Estimador da variável C. |
| \hat{C}_{NB} | Valor alvo mais provável. |
| N | Tamanho populacional. |
| n | Tamanho amostral. |
| D | Dimensão de dados. |
| Ω | Espaço amostral. |
| \bar{x} | Média amostral. |
| μ | Média populacional. |
| s | Desvio padrão amostral. |
| σ | Desvio padrão populacional. |

SUMÁRIO

| | | |
|--------------|--|-----------|
| 1 | INTRODUÇÃO | 28 |
| 1.1 | MOTIVAÇÃO | 30 |
| 1.2 | PROBLEMA DE PESQUISA | 32 |
| 1.3 | OBJETIVOS | 33 |
| 1.4 | METODOLOGIA CIENTÍFICA | 33 |
| 1.5 | ORGANIZAÇÃO DA TESE ADOTADA | 34 |
| 2 | APRENDIZADO EM FLUXO DE DADOS | 36 |
| 2.1 | APRENDIZADO ESTATÍSTICO | 37 |
| 2.1.1 | Classificador Naive Bayes | 39 |
| 2.1.2 | Classificador Hoeffding Tree | 40 |
| 2.2 | MUDANÇA DE CONCEITO | 42 |
| 2.2.1 | Velocidade da Mudança | 43 |
| 2.2.2 | Detectores de Única Análise | 45 |
| 2.2.3 | Detectores de Múltipla Análise | 51 |
| 2.3 | GERADORES E BASES DE DADOS REAIS | 53 |
| 2.3.1 | Geradores de Dados | 53 |
| 2.3.2 | Base de Dados Reais | 55 |
| 2.4 | APRENDIZADO ADAPTATIVO COM DETECTOR EMBUTIDO | 56 |
| 2.5 | CONSIDERAÇÕES FINAIS | 59 |
| 3 | ASPECTOS METODOLÓGICOS | 61 |
| 3.1 | ANÁLISE DE DADOS | 61 |
| 3.1.1 | Pré-Análise | 61 |
| 3.1.2 | Exploração dos resultados | 64 |
| 3.1.3 | Tratamento dos resultados, inferência e interpretação | 65 |
| 3.2 | AVALIAÇÃO EMPÍRICA | 66 |
| 3.2.1 | Estação de trabalho | 66 |
| 3.2.2 | Execução de experimentos | 66 |
| 3.2.3 | Geradores e bases de dados reais | 67 |
| 3.2.4 | Procedimentos de coleta de dados | 68 |
| 3.3 | CONSIDERAÇÕES FINAIS | 69 |
| 4 | DETECTORES DE MÚLTIPLA ANÁLISE | 70 |
| 4.1 | MÚLTIPLA DETECÇÃO | 72 |
| 4.2 | MOTIVAÇÃO | 74 |

| | | |
|----------|--|------------|
| 4.3 | DDE: DRIFT DETECTION ENSEMBLE | 75 |
| 4.4 | MSDDE - MIXED SIGNALS DRIFT DETECTION ENSEMBLE | 77 |
| 4.5 | MDDE - MATCH DRIFT DETECTION ENSEMBLE | 79 |
| 4.6 | RESULTADOS E DISCUSSÃO | 82 |
| 4.7 | CONSIDERAÇÕES FINAIS | 96 |
| 5 | DETECTORES DE ÚNICA ANÁLISE | 97 |
| 5.1 | USDD: ULTIMATELY SIMPLE CONCEPT DRIFT DETECTION | 98 |
| 5.2 | USDDX: USDD EXTENSION | 100 |
| 5.3 | RESULTADOS E DISCUSSÃO | 102 |
| 5.4 | CONSIDERAÇÕES FINAIS | 121 |
| 6 | MOAMANAGER | 122 |
| 6.1 | PRINCIPAIS FUNCIONALIDADES | 123 |
| 6.2 | IMPLEMENTAÇÃO | 129 |
| 6.3 | EXPERIÊNCIA DO USUÁRIO | 132 |
| 6.4 | CONSIDERAÇÕES FINAIS | 133 |
| 7 | CONSIDERAÇÕES FINAIS | 134 |
| 7.1 | TRABALHOS FUTUROS | 137 |
| | REFERÊNCIAS | 139 |
| | APÊNDICE A – AVALIAÇÃO DE DESEMPENHO | 149 |
| | APÊNDICE B – DETECTORES CONVENCIONAIS | 163 |
| | APÊNDICE C – COMITÊS DA LITERATURA | 179 |
| | APÊNDICE D – RESUMO DOS TESTES COM COMITÊS | 200 |
| | APÊNDICE E – RESUMO DOS TESTES COM TRADICIONAIS | 255 |

1 INTRODUÇÃO

Os trabalhos adotados como referenciais teóricos procuram oferecer uma representação justa do compromisso deste trabalho em assegurar garantias teóricas. A diversidade de assunto reflete, por um lado, a multiplicidade de problemas que podem ser explorados e, por outro lado, a incapacidade de permanecer em múltiplos problemas, focando em uma única área problemática para fazer explorações mais extensas. Com tanta diversidade, um título tão específico quanto *Otimização de Falsos Alarmes em Detecção de Mudança de Conceito* pode não ser muito justo em relação à pluralidade de conteúdos, entretanto, são apresentados os elos e relações que os conectam. Para este fim, são considerados os conceitos envolvidos em mais alguns detalhes.

O foco principal desta tese está nos algoritmos de Detecção de Mudança de Conceito (DMC) – que recebem o nome genérico de detectores (BARROS; SANTOS, 2018). Seu uso está associado a abordagens de Aprendizado Adaptativo com Detector Embutido (AADE) (GAMA et al., 2004). Nessa abordagem, o algoritmo adaptador testa e treina o modelo de Aprendizado de Máquina (do inglês, *Machine Learning*) (AM) e leva em consideração a análise do detector para manter e/ou iniciar um novo modelo de classificação, sendo muito útil em problemas que envolvem mudanças de conceito. Por fim, como a temática do trabalho envolve principalmente os campos de mineração de dados e aprendizado de máquina, a interseção entre eles são os dados. Sendo assim, será apresentada uma breve contextualização da importância dos dados em seu uso nos campos de pesquisa.

Esta tese delimita-se a problemas de classificação supervisionada. Dessa forma, são oferecidos elementos para contextualizar e explorar tais casos. Neste sentido, deste ponto em diante, ao se tratar de problemas de AM, fica subentendido que se tratam problemas de classificação, exceto nos casos em que seja explicitamente mencionado o contrário.

A classificação supervisionada em AM é embasada na teoria do aprendizado estatístico (CORTES; VAPNIK, 1995; VAPNIK; CHERVONENKIS, 1971). Consequentemente, são estabelecidas suposições para que seja possível obter as garantias teóricas. Respeitadas essas premissas, é possível obter classificadores com boa generalização. É importante destacar que essa capacidade é dada por previsões corretas de classificações de novos exemplos do mesmo domínio, sendo possível partir de casos específicos para casos gerais de modo satisfatório.

Uma parcela dos problemas de classificação em cenários de aplicações reais está relacionada aos dados. Nesse sentido, as informações utilizadas pelos métodos de AM são organizadas por meio de classes, isto é, grupos de dados que são organizados mediante prévia suposição a um dado domínio. Nesse caso, o termo domínio pode ser entendido como uma abstração ao conjunto de verdades assumidas para agrupar os conjuntos de dados em sua origem (criação). O problema da troca de domínio durante o

aprendizado é caracterizado pela expressão *Mudança de Conceito (do inglês, Concept Drift) (MC)* (WIDMER; KUBAT, 1996).

Dentre os recursos utilizados para mineração de fluxo de dados com mudanças de conceito, a ferramenta *Massive Online Analysis* (MOA) (BIFET et al., 2018) tem ganhado destaque ao longo do tempo. Além da aplicação, há também o *Framework* escrito em Java e desenvolvido na Universidade de Waikato, Nova Zelândia. Nele, é inclusa uma coleção de algoritmos de aprendizado de máquina, detecção de mudanças de conceito e recursos para avaliação de desempenho. Assim sendo, comumente é adotado para avaliar experimentalmente trabalhos na área de MC.

Gama et al. (2004) propuseram o *Drift Detection Method* (DDM) para detectar as mudanças de conceito. Segundo Gonçalves Jr. et al. (2014), Barros (2017), nessa estratégia de Gama et al. (2004), o componente responsável por rastrear a MC faz sua análise sob aspectos da distribuição de dados. Assim sendo, esse analista recebe como entrada um valor binário resultante de acerto (zero) ou erro (um) da previsão do classificador e entrega como saída a classificação do fluxo que pode ser uma DMC.

Várias investigações foram realizadas e novos detectores foram disponibilizados na literatura da área desde a publicação do DDM. Alguns desses detectores são mais populares que outros, pois foram utilizados em estudos por outros autores. Por exemplo, Baena-Garcia et al. (2006) propuseram uma otimização ao DDM para alguns casos de MC e criaram o método *Early Drift Detection Method* (EDDM). Além dessa, outras abordagens foram propostas, a exemplo do *ADaptive WINdowing* (ADWIN) (BIFET; GAVALDÀ, 2007a), *Statistical Test of Equal Proportions* (STEPD) (NISHIDA; YAMAUCHI, 2007), *EWMA for Concept Drift Detection* (ECDD) (ROSS et al., 2012), *Sequential Drift Change Detector 1* (SeqDrift1) (SAKTHITHASAN; PEARS; KOH, 2013), *Sequential Drift Change Detector 2* (SeqDrift2) (PEARS; SAKTHITHASAN; KOH, 2014) e *Detecting Volatility Shift in Data Streams* (SEED) (HUANG et al., 2014).

Outros exemplos mais recentes, considerando o primeiro ano que este trabalho começou a ser desenvolvido, são: *Hoeffding-based Drift Detection Method A-Test* (HDDMA), *Hoeffding-based Drift Detection Method W-Test* (HDDMW) (FRÍAS-BLANCO et al., 2015), *Selective Detector Ensemble (e-Detector)* (DU et al., 2014) e *Drift Detection Ensemble* (DDE) (MACIEL; SANTOS; BARROS, 2015). Considerando o segundo ano, tem-se: *Fast Hoeffding Drift Detection Method* (FHDDM) (PESARANGHADER; VIKTOR, 2016), *At Least One Detects Drift* (ALO), *At Least Half of the Detectors Detect Drift* (ALHD) e *All Detectors Detect Drift* (AD) (WOZNIK et al., 2016). No terceiro ano, foi proposto o *Reactive Drift Detection Method* (RDDM) (BARROS et al., 2017). Por fim, no quarto ano, foram publicados o *Wilcoxon Rank Sum Test Drift Detector* (WSTD) (BARROS; HIDALGO; CABRAL, 2018), *Fisher Test Drift Detector* (FTDD), *Fisher Square Drift Detector* (FSDD), *Fisher Proportions Drift Detector* (FPDD) (CABRAL; BARROS., 2018) e *Stacking Fast Hoeffding Drift Detection Method* (FHDDMS) (PESARANGHADER;

VIKTOR; PAQUET, 2018).

Vale ressaltar que, nos trabalhos Du et al. (2014), Maciel, Santos e Barros (2015), Wozniak et al. (2016) foi utilizada uma estratégia de múltipla análise de detecção. Entretanto, esse tipo de abordagem ainda é pouco explorada, dada a comparação em relação aos detectores de única análise. Em métodos de múltipla análise, há dois ou mais fluxos de controle (execução) para analisar os dados, diferente do que ocorre nos métodos de única análise que possuem apenas um fluxo de controle

1.1 MOTIVAÇÃO

AM em cenários de fluxos contínuos de dados desperta cada vez mais interesse da comunidade em geral, devido aos inúmeros desafios em aberto. Dentre os desafios existentes, é possível destacar a DMC, uma vez que o efeito da mudança de conceito na classificação torna o modelo impreciso ou obsoleto. A área de pesquisa de MC tem recebido significativa atenção nos últimos anos (boa parte dos trabalhos foi publicada nos últimos 10 anos). Apesar de a MC ser um problema de AM conhecido há mais de 50 anos (PAGE, 1954), o aumento no interesse por trabalhos direcionados ao estudo de MC é relativamente recente (pouco mais de 15 anos).

A MC em AM pode ser tratada por diferentes perspectivas, a exemplo de: (1) comitês de classificadores (BARROS; SANTOS; GONÇALVES JR., 2016; SANTOS et al., 2014; MINKU; WHITE; YAO, 2010; MINKU, 2010; MINKU; YAO, 2012; SIDHU; BHATIA, 2018; BRZEZIŃSKI; STEFANOWSKI, 2011; BRZEZIŃSKI; STEFANOWSKI, 2013; BARROS, 2017); (2) *drifts* recorrentes (GONÇALVES JR.; BARROS, 2013a); e (3) pares de classificadores (BACH; MALOOF, 2008). Apesar de existirem inúmeras abordagens para lidar com MC, este trabalho tem interesse especial na abordagem de AADE, utilizando apenas um modelo de classificação. Mais especificamente, pretende-se explorar técnicas de DMC.

Apesar do uso de comitês ainda ser pouco explorado para DMC, abordagens de comitês são bastante exploradas em problemas de classificação. Evidências empíricas e aplicações específicas de AM confirmam que um dado indutor pode superar vários outros para um problema específico ou para um subconjunto específico, mas é incomum encontrar um único especialista que obtenha os melhores resultados em muitos cenários (DIETTERICH, 2000). Como consequência, *Ensemble Methods* (métodos de comitê) procuram explorar o comportamento local de diferentes especialistas para aumentar a precisão da classificação geral. Além disso, supondo que algum especialista falhe, o sistema global pode corrigir o erro.

O uso de comitês apresenta-se como alternativa competitiva para a resolução de desafios do mundo real e, em especial, sua aplicação para detectar MC. Nessa perspectiva, dois ou mais detectores com características de análises diferentes podem ser empregados para obter uma solução mais satisfatória em termos de acurácia do modelo

e/ou detecção.

É possível considerar como recente o uso de abordagens de comitês para resolver problemas de DMC. Há diferentes questões em aberto, principalmente em relação às suas vantagens. Nesse sentido, pretende-se formular respostas para algumas dessas questões em aberto, em especial atenção ao problema de otimização de falsos alarmes de DMC. Além disso, pretende-se investigar os benefícios do uso de comitê de detectores na acurácia da classificação.

Além dos pontos apresentados, os seguintes pontos motivam esta tese para reduzir falsos alarmes de DMC utilizando comitês de detectores:

- São utilizados comitês com sucesso em problemas de classificação. O teorema da inexistência do almoço grátis *No-Free-Lunch Theorem* ou **NFL** (WOLPERT; MACREADY, 1995; WOLPERT, 1996; WOLPERT; MACREADY, 1997; HO; PEPYNE, 2002) afirma que todos os algoritmos de busca (incluindo os detectores) possuem desempenhos iguais quando é obtida a média por meio de todos os infinitos problemas existentes. Em outras palavras, não existe um classificador único que possa ser considerado ótimo para todos os cenários. Contudo, existem poucos trabalhos disponíveis na literatura da área sobre comitês de detectores. Efetivamente uma ótima oportunidade para estudo, desenvolvimento de hipóteses e constatações.
- Comitês são sistemas não lineares que conseguem manter vários membros com estados diferentes e consideram o estado particular de cada um deles para tomada de decisão do comitê. Apesar de alguns membros apresentarem comportamento menos preciso para determinada situação, eles podem ser mais precisos em novos cenários (WOLPERT, 1996). A otimização dos resultados é alcançada quando o sistema é capaz de usar os membros mais precisos para de alguma forma melhorar o desempenho global. Investigar técnicas e estratégias capazes de maximizar os resultados individuais quando utilizados em conjunto para tomadas de decisão é um desafio ainda maior em DMC.
- Algoritmos para fluxos contínuos de dados são projetados sob restrições de uso de recursos, principalmente tempo de processamento e memória principal. Abordagens de comitês frequentemente permitem a implementação paralela, que pode ser útil para reduzir o seu tempo de execução, tornando praticamente indiferente em relação às abordagens lineares. Sendo assim, entende-se que utilizar mais de um detector pode resultar em melhores chances de obter-se detecções mais fidedignas à realidade (minimizando falsos alarmes), sem depreciar muito o fator tempo. Por fim, é importante deixar claro que soluções que envolvem comitês, geralmente possuem um custo computacional maior quando comparado a soluções convencionais.

- Reduzir sobreajuste dos dados, pois é esperado que cada detector possa convergir para distintos mínimos locais. Em casos de comitês de classificadores, usualmente são produzidos melhores resultados que um único classificador (DIETTERICH, 2000).

Mesmo sabendo que os comitês apresentaram ótimos resultados em problemas de classificação, há poucos ou quase nenhum estudo que afirme que sua utilização em problemas de mudança de conceito possam ser úteis. Investigações sobre comitês de detectores podem revelar o potencial dessa abordagem para DMC e inspirar novas investigações.

Tal estudo deve ser realizado nesta tese com o intuito de inspirar a criação de abordagens novas e eficientes para lidar com falsos alarmes em MC. Faz parte da motivação deste trabalho otimizar resultados de DMC para minimizar falsos alarmes.

Por fim, mas não menos importante, há também desafios relacionados ao uso em larga escala de experimentos na ferramenta MOA. Apesar de ser uma utilizada solução para mineração de fluxo de dados com MC, não há muitos recursos para gerenciar experimentos com muitos dados. Exemplos nesse sentido incluem automatizar tarefas repetidas ou identificar e corrigir erros de execução. Nesse sentido, também faz parte da motivação deste trabalho promover avanços à área de MC com solução de *software* capaz de otimizar o uso da ferramenta MOA.

1.2 PROBLEMA DE PESQUISA

Diante do exposto até aqui, esta tese propõe-se a responder o seguinte problema de pesquisa: **Como otimizar aspectos de falsos alarmes em detectores?**

Falso alarme é um termo genérico utilizado para mencionar *False Positive* (traduzindo para o português, Falso Positivo) (FP), *False Negative* (traduzindo para o português, Falso Negativo) (FN), *True Positive* (traduzindo para o português, Verdadeiro Positivo) (TP) e *True Negative* (traduzindo para o português, Verdadeiro Negativo) (TN). Essas métricas são adotadas para estimar um compromisso justo entre os reais pontos de mudanças de conceito e as estimativas de mudanças de conceito realizadas pelos métodos. As métricas fornecem importantes indicações em relação ao quantitativo correto e incorreto dos reais pontos de mudanças e detecções realizadas. As métricas de FP e FN sem o complemento das métricas de TP e TN, não são completas para a avaliação desejada dos falsos alarmes de detecção, pois não fornecem indicadores da real quantidade de mudanças de conceito.

Segundo (BIFET et al., 2013), a qualidade de um detector pode ser dada por um compromisso justo entre rápidas detecções e erros de estimação (detecções erradas). Segundo (BIFET, 2017), é comum na literatura da área a avaliação detectores usando acurácia *Prequential* da classificação, entretanto, essa avaliação não representa um compromisso justo em relação ao desempenho do detector. Assim sendo, deseja-se avaliar esse compromisso utilizando a métrica de *Matthews Correlation Coefficient*

(traduzindo para o português, Coeficiente de Correlação de Matthews) (MCC) para avaliar o desempenho dos métodos de detecção propostos em relação a otimização dos resultados.

1.3 OBJETIVOS

O principal objetivo desta tese é responder ao problema de pesquisa e suas questões relacionadas. O assunto tratado é sobre detecção de mudança de conceito e o tema central desta tese está na discussão da **otimização de falsos alarmes de detecção**. Para isso, são estudados alguns aspectos do problema (assuntos relacionados) da detecção de mudanças de conceito.

Nesta tese são desenvolvidos métodos de DMC, fazendo parte do objetivo mapear seu estado da arte. A seguir são apresentados os objetivos específicos:

- investigar as principais noções conceituais sobre técnicas de DMC;
- desenvolver *softwares* de otimização para a ferramenta MOA;
- melhorar o estado da arte em DMC;
- desenvolver métodos de DMC de múltipla e única análise.

1.4 METODOLOGIA CIENTÍFICA

De acordo com Prodanov e Freitas (2013), esta tese classifica-se em relação ao tipo de pesquisa como:

- **Quanto à natureza (finalidade da pesquisa):**
 - Aplicada, pois o objetivo é gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos. Envolve verdades e interesses locais.
- **Quanto à forma de abordagem do problema:**
 - Qualitativa e Quantitativa.
- **Quanto aos fins da pesquisa:**
 - Descritiva, pois deriva de natureza aplicada.

Em relação às características do tipo de pesquisa:

- **Quanto aos procedimentos:**
 - Experimental, pois deriva de natureza aplicada e fins de pesquisa descritiva;
- **Gerais:**
 - Verifica-se a relação entre variáveis, necessário para procedimento experimental.
- **Tipos de instrumento:**
 - Plano de pesquisa, utilizado para manipular condições e observar os efeitos produzidos, necessário para procedimento experimental.

1.5 ORGANIZAÇÃO DA TESE ADOTADA

Objetivando uma leitura menos árida e fria, foram adotadas algumas estratégias. Nesse sentido, para cada seção desta tese, foi adotada a estratégia de:

- Começar com um resumo da seção;
- Posteriormente, são apresentadas considerações iniciais a serem introduzidas na seção;
- Quando for julgado pertinente, estabelecer um paralelo da teoria com tais situações do mundo real (situação concreta);
- Evitar desenvolvimentos de natureza puramente abstrata (tais como demonstrações matemáticas), porém mantendo o nível de rigor adequado para uma tese, dado o conteúdo a ser apresentado;
- Complementar a teoria com exemplos, relativos a diferentes áreas de aplicações, bem como a questões da vida cotidiana;
- Concluir com as considerações do autor em relação ao conteúdo apresentado na seção.

Pode-se entender que há vários problemas no mundo real que exigem soluções. Entretanto, antes de trabalhar nas soluções, é preciso pensar no problema. Assim sendo, problemas podem ser agrupados, por exemplo, identificando ou descrevendo características que causam determinada situação ou por classificações que agrupam essas características. Logo, pode-se levar em consideração, soluções para cada tipo de problema, sendo comum adotar premissas para validade da solução. Dessa forma, o reconhecimento de que existem vários tipos de problemas é importante para possibilitar a escolha da solução mais apropriado para cada tipo situação/problema.

Uma breve descrição das seções deste trabalho de doutorado é apresentada a seguir:

- A Seção 2 analisa a literatura sobre MC, aprendizado em fluxo de dados com MC, além de geradores e conjuntos de dados do mundo real utilizados nos experimentos desta tese. Além disso, também são mostrados os principais trabalhos relacionados. Há uma complementação no Apêndice A, que trata avaliação de desempenho e principais métricas adotadas;
- A Seção 3 descreve em detalhes os procedimentos utilizados nos experimentos, parametrização de métodos, geradores de dados, cenários, abordagens estatísticas para análise e comparações. Procurou-se reunir as informações relevantes aos procedimentos experimentais e à análise dos resultados da tese em uma única seção;
- As Seções 4 e 5 descrevem em detalhes os métodos propostos. Adicionalmente, são apresentados os resultados experimentais em relação aos trabalhos para

identificação dos melhores desempenhos em termos de acurácia, *Mean Time Ratio* (traduzindo para o português, Taxa Média de Tempo) (MTR) e MCC na avaliação de desempenho dos trabalhos relacionados, incluídos nos apêndices B e C. Além disso, nos apêndices D e E são apresentados os resultados experimentais utilizados para avaliação dos métodos propostos em relação aos trabalhos relacionados. Adicionalmente, o algoritmo do DDE é apresentado com algumas melhorias, originalmente ele foi publicado na conferência *27th IEEE International Conference on Tools with Artificial Intelligence* (ICTAI), na Itália em 2015;

- A Seção 6 descreve a ferramenta *Massive Online Analysis Manager PHP framework* (MOAManager) em detalhes, incluindo sua motivação, principais funcionalidades e conclusões. Ela é baseada no conteúdo do artigo *MOAManager: a tool to support data stream experiments*, recentemente publicado (MACIEL; SANTOS; BARROS, 2020);
- A Seção 7 descreve as considerações finais e apresenta propostas de trabalhos futuros.

2 APRENDIZADO EM FLUXO DE DADOS

Esta seção 2 apresenta conceitos, técnicas e desafios do aprendizado em ambientes de fluxo de dados com mudanças de conceito. Para isso, inicialmente são apresentados problemas decorrentes da qualidade da modelagem de dados. Em seguida são apresentados desafios para realizar aprendizado mediante alguns problemas inerentes à área de aprendizado de máquina. Por fim, são apresentadas abordagens para minimizar os efeitos da aparição de MC em fluxos de dados.

Em aprendizado de máquina, a qualidade dos dados na classificação pode ser comprometida por diferentes problemas, incluindo ruído, classes desbalanceadas, dados complexos, valores ausentes, elevado número de atributos (maldição da dimensionalidade), erros nas regras de geração dos dados e, como é o caso desta tese, a mudança nas regras de classificação dos dados ao longo do tempo. Essa mudança nas regras de classificação ocorre quando o conjunto de regras é alterado de modo imprevisível, durante ou depois do conceito em aprendizado ser estabilizado, afetando significativamente a capacidade do modelo para generalizar as hipóteses corretamente (GAMA et al., 2014; KRAWCZYK et al., 2017). Em cenários do mundo real, essas mudanças são mais comuns, em função da flexibilidade e do dinamismo atrelados ao ambiente, implicando classificações que não pertençam às mesmas distribuições (JAPKOWICZ, 2016).

Independente do modo (com ou sem supervisão) e/ou forma (incremental ou não) de aprendizado adotado, mudar o conjunto de regras que geram as classificações dos dados prejudica a construção e/ou evolução do modelo significativamente. Esse fenômeno da mudança é chamado de *mudança de conceito* (do inglês, *Concept Drift*). Pode-se concluir que MC refere-se a uma mudança no conjunto de regras que geram os dados ao longo do tempo, e essas regras dão origem a um conceito, isto é, uma distribuição de dados (em outras palavras, o conceito a ser aprendido) (GAMA et al., 2014; WIDMER; KUBAT, 1996). Portanto, essas definições apresentadas convergem para descrever ambientes não estacionários, logo, é adequado adotar uma definição probabilística para descrever MC (ELWELL; POLIKAR, 2011).

Um dos desafios em AM é a aprendizagem de conceito. Pode-se entender que conceitos são categorias compostas por associações que ajudam a classificar alguma coisa (física ou abstrata), construída no entendimento de que cada classificação possui um conjunto de características relevantes e comuns. Assim, o aprendizado de conceito é uma estratégia que exige comparações e dessemelhança entre duas coisas ou categorias.

Em AM, no paradigma supervisionado, um indutor toma conhecimento de um fato obtendo informações (por experiências). Normalmente, o indutor observa uma determinada quantidade de exemplos para tornar-se capaz de generalizar casos, mesmo

para casos sem exemplo, conseguindo estimar a classificação de novos casos. Contudo, a aprendizagem de conceito pode depender de um contexto não presente no conjunto de fatos colhidos. Logo, mudanças nesse contexto podem fazer com que ocorram mudanças no conceito a ser interpretado pelo indutor.

Deste modo, um indutor criado para realizar classificação em ambiente de fluxos de dados precisa lidar com contextos ocultos. Uma mudança de contexto pode comprometer o conceito aprendido dos fatos colhidos, pois esses fatos são utilizados para prever a classificação de novos exemplos. Nesse cenário, a abordagem de classificação precisa ser robusta para diferenciar dados ruidosos e mudança de conceito, evitando descartar indevidamente o conceito aprendido.

Até o presente momento o conceito de AM foi amplamente explorado e, a partir daqui, será mais discutido o problema da DMC. Nas próximas subseções são fornecidos os princípios gerais de Detecção de Mudança de Conceito e uma visão geral do AADE.

2.1 APRENDIZADO ESTATÍSTICO

A teoria do aprendizado estatístico surgiu por volta da década de 60 (VAPNIK; CHERVONENKIS, 1971; CORTES; VAPNIK, 1995; VAPNIK; CHERVONENKIS, 2015). Por volta de 1990, surgiram novas propostas e tipos de algoritmos de aprendizado. Os avanços tecnológicos e no campo de aprendizado estatístico fizeram com que os conceitos teóricos propostos ao longo dos anos pudessem ganhar vida e serem aplicados na construção de algoritmos práticos para estimar funções.

Existe uma gama de algoritmos para abordagens de classificação. Eles classificam cada exemplo para uma dada distribuição de probabilidade dentro do espaço total de probabilidades, assim determinando um conjunto limite de decisões de acordo com o processo de supervisão do algoritmo. Um classificador pode ser denotado como uma função $f : x \rightarrow y$, no qual x é um conjunto de entradas da função, e y uma saída pertencente ao conjunto de saídas. Logo, f representa o modelo, isto é, a função de classificação.

O conjunto de entradas x é formado por atributos que fazem referência à tarefa de classificação. Esses atributos precisam representar bem o conjunto de características para que haja o aprendizado do conceito de modo satisfatório. O conjunto de saídas é formado por todas as classes que serão previstas e aprendidas pelo modelo, dados os atributos do conjunto x , pertencentes ao seu exemplo.

Uma das motivações para se estudar a teoria do aprendizado estatístico é para provar que é possível um algoritmo aprender a partir de exemplos, assim como saber o quão bom esse algoritmo é. O aprendizado teórico tem como objetivo oferecer garantias teóricas para que esses algoritmos quando possuem dados suficientes (o quantitativo de dados é relativo ao problema de classificação), obtenham uma alta taxa de previsões corretas. Logo, há garantias teóricas de aprendizado a partir de uma boa taxa de classificação correta dada

uma certa quantidade de dados.

A teoria do aprendizado estatístico possui algumas premissas para que haja garantias teóricas em relação ao aprendizado (CORTES; VAPNIK, 1995). Em síntese, essas premissas são:

1. Nenhuma suposição é feita sobre a função de probabilidade conjunta (qualquer distribuição conjunta de $Pr(x,y)$ pode ser considerada). Uma Distribuição conjunta $Pr(x,y)$ significa que dadas entradas/exemplos em x , e saídas/rótulos em y , ela é construída dentro de um espaço de possibilidades. Cada possível resultado tem seu valor de probabilidade associado, no caso, se tiver todas as frequências dos possíveis valores em x e possíveis saídas em y . O conjunto (os valores combinados) mais provável será o que tem a maior probabilidade associada. É possível se ter qualquer distribuição porque ela pode abranger tanto gaussianas como outras;
2. Dados amostrados de forma igualmente distribuídas. Isso significa que é possível amostrar exemplos de qualquer distribuição de probabilidade. O primeiro exemplo foi classificado como sendo $Pr(1,0)$, o segundo não será necessariamente da mesma distribuição de probabilidade $Pr(1,0)$; ele pode ser de $Pr(0,1)$, $Pr(0,0)$, ou da própria $Pr(1,0)$;
3. Exemplos independentes entre si. Isso significa que não é possível induzir que o próximo exemplo seja classificado dentro de uma mesma distribuição de probabilidade, ou qualquer outra imposta. Ser independente é ter uma amostra em qualquer região do espaço de amostras. No caso, a probabilidade de amostragem dos exemplos não muda, mesmo que seja selecionada a mesma distribuição mais de uma vez, ou seja, a probabilidade do evento ocorre não muda ao longo de novos exemplos. Se as probabilidades mudassem, significaria que os exemplos seriam dependentes;
4. Os rótulos/classes podem ter ruídos. É possível que algumas classificações tenham sido feitas de forma errada, e será passado para aprender aquele exemplo de forma incorreta. O objetivo é que a aprendizagem seja capaz de identificar que houve erro naquela classificação considerada como correta - obviamente isso só é possível se esses exemplos classificados incorretamente não forem a maioria (não pode dominar a distribuição de probabilidade);
5. A distribuição de probabilidade $Pr(x,y)$ tem que ser estática, isso significa dizer que, a distribuição não pode mudar ao longo do tempo. O motivo é que, se elas mudam, não é possível ter uma boa estimativa de acordo com o que já foi coletado. Mesmo para quem trabalha com fluxo de dados, é necessário se ter uma distribuição estática para garantir o aprendizado. Assim sendo, é importante destacar que essa

distribuição estática é necessária para que seja possível conquistar uma convergência do modelo, mas são exploradas técnicas para adaptar o modelo a cenários com mudanças de conceito;

6. A distribuição $Pr(x,y)$ é supostamente desconhecida durante a fase de treinamento. Só é conhecida uma amostra dessa distribuição. Dados são coletados para estimar a distribuição. Se todas as distribuições fossem conhecidas, não precisaria haver estimativas (y).

Por fim, a teoria do aprendizado estatístico (CORTES; VAPNIK, 1995) oferece garantias teóricas para conjuntos de dados estacionários. Assim sendo, uma mudança de conceito durante o aprendizado torna o modelo obsoleto em termos teóricos. Isso ocorre porque não haverá mais garantias de minimização do erro de classificação.

2.1.1 Classificador Naive Bayes

Segundo Lewis (1998), Gama e Castillo (2002, p.1), é comum em trabalhos de AM adotar um classificador Bayesiano, também conhecido pelo termo *Naive Bayes*. Esse termo foi utilizado pela primeira vez por Richard Duda e Peter Hart na obra intitulada “*Pattern Classification and Scene Analysis*”, em 1973 (DUDA; HART, 1973). Sua natureza de classificação é probabilística e fundamenta-se no teorema de *Bayes*, sendo denominado como “ingênuo” (HAND; YU, 2001) (do inglês, Naive) por assumir que os atributos são condicionalmente independentes, ou seja, a informação de um evento não é dependente de outro (KOTSIANTIS, 2007, p.257-258). Apesar do pressuposto da independência, na prática, ele funciona surpreendentemente bem, mesmo quando a hipótese de independência não é verdadeira (RUSSELL; NORVIG, 2013, p.587).

Um das principais vantagens do *Naive Bayes* é sua simplicidade para formular as hipóteses, velocidade e baixo uso de recursos computacionais (DUDA; HART; STORK, 2001). Além disso, como o classificador utiliza a regra do produto, o número pode ser convertido em logaritmos grandes, implicando vantagens computacionais significativas (KOTSIANTIS, 2007, p.257)(JOHN; LANGLEY, 1995; BIFET; FRANK, 2010). A Regra do produto pode ser expressa de duas formas, definidas nas equações 2.1 e 2.2:

$$Pr(x \wedge y) = Pr(x | y) * Pr(y) \quad (2.1)$$

$$Pr(x \wedge y) = Pr(y | x) * Pr(x) \quad (2.2)$$

Igualando os dois membros da equação 2.1 e dividindo por $Pr(x)$, pode ser reescrita de outra maneira, a saber:

$$Pr(y | x) = \frac{Pr(x | y) * Pr(y)}{Pr(x)} \quad (2.3)$$

Essa equação 2.3 é conhecida como regra de *Bayes* (e também como lei de *Bayes* ou teorema de *Bayes*). Essa equação é a base de todos os sistemas modernos de inteligência artificial para inferência probabilística (RUSSELL; NORVIG, 2013). Os *Naive Bayes* classificam seu modelo com uma regra de decisão. A regra comum é escolher a hipótese mais provável. Isso é conhecido como estimativa máxima de probabilidade a posteriori (do inglês, *maximum a posteriori probability* - MAP) e usado como regra de decisão. No *Naive Bayes* (NB) aplica-se a regra MAP na equação 2.4.

$$C_{MAP} = \arg \max_{c \in C} \prod_{i=1}^n Pr(X|c_i)Pr(c_i) \quad (2.4)$$

Como o NB lida com o processo de indução do modelo de classificação do conjunto de dados rotulados, ele utiliza uma variação da regra de *Bayes* (equação 2.3) para prever a classe para uma instância de teste, assumindo que as características são condicionalmente independentes umas das outras, dada a classe. A equação 2.5 descreve o classificador NB.

$$\hat{C}_{NB} = \arg \max_{c_j \in C} Pr(c_j) \prod_{i=1} Pr(a_i|c_j) \quad (2.5)$$

Na equação 2.5, dada uma instância, o classificador \hat{C}_{NB} recebe uma classe (**argmax**, isto é, a classe c com máxima probabilidade é escolhida para a instância) da função objetivo

$$Pr(c_j) \prod_{i=1} Pr(a_i|c_j) \quad (2.6)$$

2.1.2 Classificador Hoeffding Tree

O algoritmo *Hoeffding Tree* utilizado nesta tese é uma árvore de decisão incremental para fluxos de dados, entretanto, pressupõe-se que a distribuição de dados não muda ao longo do tempo. Em especial, foi utilizado a variação do algoritmo *Hoeffding Tree* (HT) disponibilizado em Manapragada, Webb e Salehi (2018) que também disponibiliza uma variação do HT chamada de *Hoeffding Anytime Tree* (HATT) que possui pequenas modificações em relação ao HT. Vale ressaltar que, os dois algoritmos estão no repositório oficial do projeto MOA. O objetivo do HATT é tornar o método mais eficiente em termos de complexidade de tempo e acurácia do modelo. Apesar de haver outras abordagens baseadas em ou variantes que são preparadas para mudanças de conceito como *Concept-adapting Very Fast Decision Tree learner* (CVFDT) (HULTEN; SPENCER; DOMINGOS, 2001), *Hoeffding Adaptive Tree* (HAT) (BIFET; GAVALDÀ, 2009) e *Adaptive Random Forest* (ARF) (GOMES et al., 2019).

O HT expande a árvore de decisão de forma incremental com base nas garantias teóricas do limite de Hoeffding (também conhecido como limite de Chernoff). Um nó é expandido assim que houver evidência estatística suficiente de que existe um recurso

de divisão ideal, uma decisão baseada no limite Hoeffding independente de distribuição. O modelo aprendido pelo *Hoeffding Tree* é assintoticamente idêntico àquele construído por um indutor não incremental, se o número de instâncias de treinamento for grande o suficiente.

Uma árvore de decisão é uma estrutura hierárquica composta por nós e *links*. O primeiro nó é chamado raiz e, por convenção, se localiza no topo da árvore. A partir do nó raiz, se estendem ramificações sucessivas denominadas *links* ou ramos que por sua vez, conectam os demais nós. Os últimos nós da árvore são denominados terminais ou folhas e representam uma decisão a ser tomada (MITCHELL, 1997).

HT é um algoritmo de classificação incremental popular para lidar com as mudanças de conceito. Suporta a indução de árvores de decisão e fornece soluções para os seguintes desafios:

- Há incerteza no tempo de aprendizagem. No HT, o tempo de aprendizado é constante por exemplo (instância) e isso significa que o classificador é adequado para a mineração de fluxos de dados;
- As árvores resultantes são quase idênticas às árvores construídas pela aprendizagem em lote convencional (MARRÓN et al., 2016), desde que receba instâncias suficientes para treinar e construir as árvores.

HT assume que a geração de distribuição dos exemplos não é constante e explora o fato de que uma pequena amostra pode ser suficiente para escolher um atributo com boa separação entre as classes, o que é matematicamente suportado pelo conceito de Hoeffding bound (HOEFFDING, 1963; MARON; MOORE, 1993). Este conceito afirma que, com probabilidade $1 - \delta$, a verdadeira média da variável é de pelo menos $\bar{r} - \varepsilon$, onde \bar{r} é o valor médio calculado a partir de n observações independentes, cujo intervalo é \mathbb{R} (por exemplo, dado uma probabilidade com intervalo 1, e, para um ganho de informação, o intervalo é $\log c$, onde c é o número de classes) e ε está determinado pela equação 2.7.

$$\varepsilon = \sqrt{\frac{R^2(\ln 1/\delta)}{2n}} \quad (2.7)$$

Hoeffding bound quantifica o número de observações quando é necessário estimar o quão bom um atributo é (DOMINGOS; HULTEN, 2000). O que torna a *Hoeffding bound* atrativo é a sua capacidade de dar os mesmos resultados independentemente da distribuição de probabilidade, gerando, assim, as observações. No entanto, o número de observações necessário para atingir certos valores de δ e ε é diferente entre as distribuições de probabilidade. Vale ressaltar que, apesar dos pressupostos por trás da aplicação do *Hoeffding bound* não estarem corretos, estas árvores continuam apresentando resultados interessantes (MATUSZYK; KREMPL; SPILIOPOULOU, 2013).

2.2 MUDANÇA DE CONCEITO

Um conceito é formado por um domínio de verdades (regras), no qual assume-se como verdade absoluta para transformar dados brutos em informação. Uma regra dentro da modelagem de um domínio pode ser abstraída como entendimento sobre fatos. Uma mudança no conjunto de regras ao longo do tempo pode ter severas implicações, entretanto, algumas delas são inevitáveis por conta da necessidade de evolução/adequação do sistema. Por fim, MC é o termo empregado para qualquer mudança nas regras que altera o entendimento sobre os fatos durante uso da modelagem.

Nesse sentido, o termo modelagem pode ser entendido como uma técnica usada para a especificação das regras de negócios. O modelo é um esquema que representa um tipo de categoria, tratado pelo termo *classe* em problemas de classificação. O produto da modelagem é um modelo abstrato de dados que tem por finalidade descrever de maneira conceitual os dados que pertencem a um domínio. Logo, modelar significa construir um modelo que explique as características dos dados, permitindo seu entendimento por meio de principais características.

Definição 1 – *Mudança de Conceito* – Em ambientes dinâmicos com variáveis não estacionárias, uma mudança de conceito pode ser vista como uma distribuição de dados que pode mudar ao longo do tempo, causada por características insuficientes, desconhecidas ou não observáveis em um conjunto de dados, produzida pelo fenômeno conhecido como contexto oculto (do inglês, *hidden context*) (WIDMER; KUBAT, 1996; SCHLIMMER; GRANGER, 1986).

Definição 2 – *Mudança de Conceito* – Uma mudança de conceito real refere-se a mudança na distribuição condicional de saída (regra de classificação dos atributos) dada uma entrada (os atributos), podendo a distribuição de entrada permanecer inalterada (GAMA et al., 2014).

O termo conceito foi introduzido no texto pela expressão *domínio de verdades*. Assim sendo, possuem significados equivalentes. A mudança em si é chamada de *mudança de conceito*, levando em consideração as definições 1 e 2, pode-se concluir que MC refere-se a mudanças no conjunto de regras que geram os dados ao longo do tempo, e essas mudanças dão origem a um conceito. No sentido estatístico, diz-se distribuição de probabilidade de dados (função de densidade), conceitualmente expresso na definição 2, dando à MC um sentido probabilístico. Portanto, essas definições apresentadas convergem para descrever fluxos de dados não estacionários. Logo, é adequado adotar uma definição probabilística para descrever MC (ELWELL; POLIKAR, 2011).

Webb et al. (2016) fazem uma caracterização das mudanças de conceito e como resultado apresentam uma taxonomia de categorias de mudanças. Assim sendo, é

possível caracterizar uma MC de diferentes modos. Entretanto, para o escopo desta tese os modos são classificados em relação à velocidade e em relação aos padrões da mudança. Na próxima seção, será abordada a classificação quanto à velocidade da mudança.

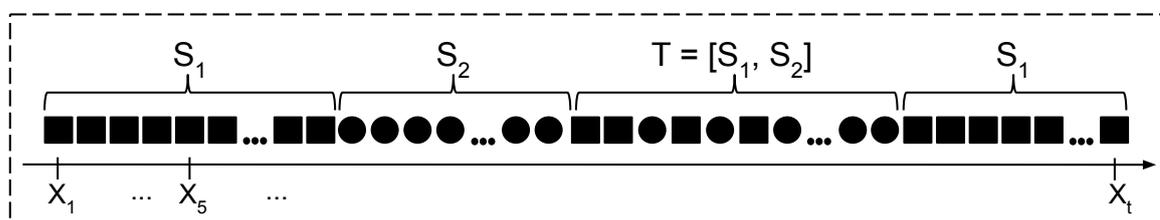
2.2.1 Velocidade da Mudança

A MC pode manifestar-se de duas maneiras: abrupta ou gradual (BARROS; SANTOS, 2019; GAMA et al., 2014; ŽLIOBAITĖ, 2009; MINKU; WHITE; YAO, 2010). A classificação entre elas é definida de acordo com a velocidade como ocorre a mudança, e essa velocidade é caracterizada pela relação de quantidade de informação pertencente ao novo conceito presente no fluxo de dados. Em mudanças abruptas, a transição entre conceitos é rápida, pois há uma substituição imediata de 100% do conceito antigo (MINKU; WHITE; YAO, 2010); já nos casos de mudanças graduais, ocorre uma mistura entre os dois conceitos durante a transição até que o novo conceito manifeste-se de modo onipresente. Por fim, as transições em mudanças graduais podem ser lentas ou rápidas, variando exclusivamente pela velocidade da mudança.

De acordo com Minku, White e Yao (2010), a velocidade da mudança é inversamente proporcional ao tempo de mudança. Ela é expressa pelo número de informações (quantidade de passos) necessárias para o novo conceito sobrescrever completamente o antigo. Logo, uma alta velocidade possui uma quantidade menor de informações.

A Figura 1 ilustra casos típicos de mudanças de conceito. Sejam S_1 e S_2 conceitos, entre S_1 e S_2 ocorre uma MC abrupta, pois houve uma substituição de S_1 por S_2 . Entre S_2 e S_1 ocorre uma MC gradual, pois houve um momento de transição T no qual os conceitos estiverem misturados.

Figura 1 – Exemplo de mudança de conceito. Seja S_1 e S_2 conceitos, T é dado pela mistura entre eles e X_1, \dots, X_t as microinformações do domínio do conceito.



Fonte: O AUTOR (2020)., adaptado de Žliobaitė (2009)

É imprevisível o modo como ocorre a mistura entre conceitos. Do ponto de vista probabilístico de conceito, durante uma transição as probabilidades que condicionam o aparecimento de cada conceito é variável, por exemplo, pode haver situações e/ou momentos nos quais há uma maior probabilidade para um dos conceitos. O tratamento dessa mistura entre conceitos pode influenciar na interpretação das microinformações,

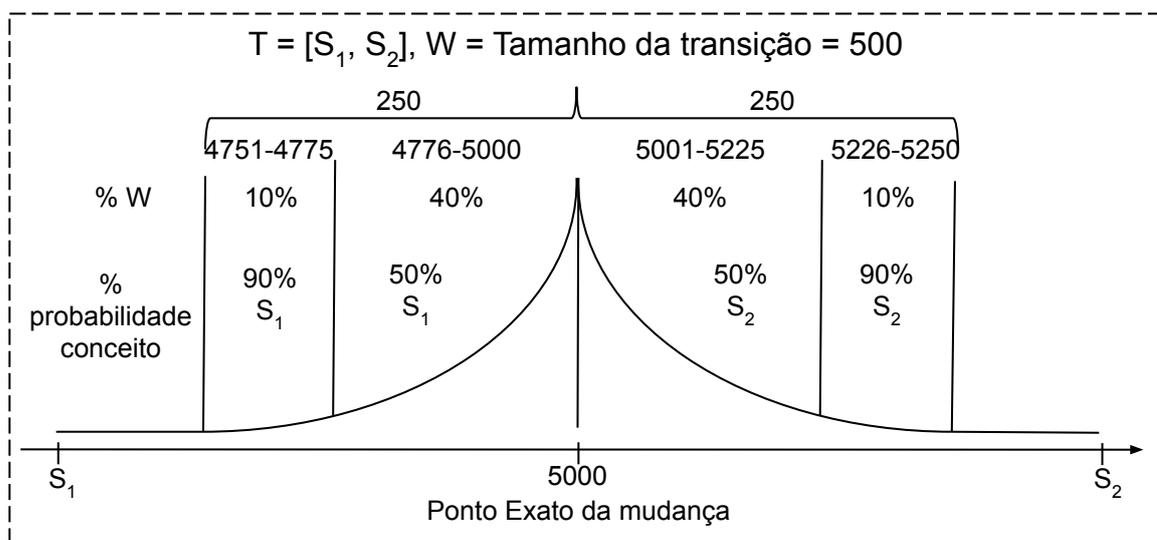
assim sendo, mudanças de conceito graduais são mais difíceis de serem detectadas e, dependendo da semelhança entre os conceitos, podem se tornar muito mais difíceis.

O *Framework* MOA utilizado para realizar simulações de mudanças de conceito implementa as mudanças graduais de acordo com a Figura 2. Segundo Barros e Santos (2019) e repositório do MOA, a classe ***ConceptDriftStream.java*** dentro do pacote ***moa.streams*** é responsável por implementar a política de transição. Por padrão, essa classe utiliza as proporções para mistura entre conceitos apresentados na Figura 2.

Ainda de acordo com a Figura 2, o valor de w define o tamanho/velocidade da transição. Seja $p=5000$ o ponto exato da MC, a proporção de w entre os conceitos misturados é expressa por $g = w * 1/2$ (ou seja, 50% do tamanho de w). A mistura entre conceitos tem início antes de p , precisamente calculado expresso por $TS_1 = p - g$ e após p por $TS_2 = p + g$. Ao misturar S_1 e S_2 , são considerados três períodos com probabilidades que variam, expressos na Figura 2. Por exemplo, nos primeiros 10% de w , S_1 tem 90% de probabilidade de ocorrer ao comparar com S_2 que terá 10%. Posteriormente, nos próximos 80% de w , S_1 e S_2 igual percentual de probabilidade de 50%. No último trecho, ou seja, que corresponde aos percentuais de 90% a 100% de w , S_2 que terá 90% de probabilidade enquanto S_1 que terá apenas 10%.

Vale ressaltar que, durante a graduação, isto é, durante a transição da MC, não é possível oferecer garantias teóricas em relação ao domínio dos dados. Implicações teóricas desse tipo de MC serão abordadas na seção aprendizado de máquina em fluxo de dados com MC.

Figura 2 – Exemplo de mudança de conceito implementada no framework MOA. Seja S_1 e S_2 conceitos, T é dado pelo período de transição entre conceitos. A mistura entre conceitos, expresso por T , corresponde ao período no qual há dois conceitos S_1 e S_2 , ou seja, duas distribuições de variáveis.



Fonte: O AUTOR (2020).

2.2.2 Detectores de Única Análise

Os trabalhos relacionados de única análise são apresentados e descritos aqui.

O método DDM tem sua estratégia baseada na análise de desempenho do classificador (GAMA et al., 2004). Seu critério de decisão é de acordo com a oscilação da taxa de erro da classificação, pois é considerado como premissa que essa taxa de erro diminui ou será constante à medida que o número de observações aumente. Portanto, a classificação do estado do fluxo de dados ocorre por conta da oscilação da taxa de erro e limites previamente definidos para *warning* e *drift*.

A cada iteração três variáveis de controle são atualizadas, m_p , m_s e m_n . Considere m_p como a taxa (média) de erro da classificação calculado por

$$\frac{m_p + (prediction - m_p)}{m_n},$$

sendo $prediction = \{0, 1\}$ que representa *sucesso* = [0] ou *insucesso* = [1] da previsão do classificador e m_s é calculada pelo desvio padrão, expressa por

$$\sqrt{\frac{m_p \times (1 - m_p)}{m_n}},$$

em relação ao número de observações m_n . A soma das duas variáveis, m_p e m_s representa o sinal de oscilação utilizado para avaliar a condição atual do fluxo de dados. Entretanto, a avaliação do sinal apenas acontece depois que um número mínimo de observações for alcançado.

Sempre que o sinal de oscilação ($sinal = m_p + m_s$) é menor ou igual a m_{psmin} , três variáveis são atualizadas, $m_{pmin} = m_p$, $m_{smin} = m_s$, $m_{psmin} = m_p + m_s$. Essa condição ocorre para controlar o intervalo de valores mínimos.

Em relação aos níveis de detecções, eles são identificados pelo sinal de oscilação. Cada um dos três possíveis estados são classificados de acordo com o sinal, no qual a primeira condição a ser verificada determina se ocorreu um *drift*, expressa por

$$sinal > m_{pmin} + (3 \times m_{smin}),$$

e a condição de *warning* é caracterizada por

$$sinal > m_{pmin} + (2 \times m_{smin}),$$

caso nenhuma das duas condições seja verdadeira, o fluxo é classificado como estável.

Sempre que o sinal ultrapassa o limite definido para o nível de *drift*, as variáveis são reinicializadas. Propositamente, m_{pmin} , m_{smin} e m_{psmin} são inicializadas com um número real de alto valor. As variáveis m_p e m_n são inicializadas com 1 e m_s com zero.

O EDDM (BAENA-GARCIA et al., 2006) é semelhante ao DDM. Ambos os métodos foram implementados seguindo a mesma particularidade básica, exceto que, ao contrário

do DDM, o EDDM foi criado com o objetivo de melhorar as detecções dos *drifts* graduais, considerando a distância entre os erros de classificação ao invés da média do número de erros. Dessa forma, é evidente que a distância entre os erros de predição do classificador aumenta à medida que o algoritmo de detecção assume que a distribuição aprendida mantém o comportamento estável. No caso contrário a esta teoria, o método leva em conta que houve uma mudança de conceito na distribuição dos dados.

Ao contrário do DDM, o EDDM calcula a distância média entre os erros (p_i') e o seu desvio padrão (s_i'). As variáveis p_{max}' e s_{max}' também são mantidas sendo propositalmente inicializadas com um número real de baixo valor. Na chegada de cada novo erro do classificador, o método estabelece a comparação das expressões seguintes: $(p_i' + 2 \times s_i') > (p_{max}' + 2 \times s_{max}')$. No caso em que a comparação anterior seja verdadeira, p_{max}' recebe o valor de p_i' e para s_{max}' é atribuído s_i' .

O ADWIN (BIFET; GAVALDÀ, 2007a) é baseado no uso da estrutura de dados tipo fila circular dinâmica, comumente chamada na literatura de área de janela deslizante. Essa janela recebe o nome de w e possui valores de desempenho com um tamanho variável ou dinamicamente ajustável. Assim, na medida que os dados lidos do processo de inferência continuarem dentro de uma determinada distribuição de probabilidade, o tamanho de w vai ser maior. No caso que o tamanho de w seja menor, o método vai detectar a presença de mudanças de conceitos. Além disso, uma vez que o preditor fornece os erros e acertos na distribuição, o ADWIN divide w em duas sub-janelas (w_0 e w_1) e realiza a comparação entre as suas médias aritméticas ($\hat{\mu}_{w_0}$ e $\hat{\mu}_{w_1}$). Dessa forma, a janela w terá o seu tamanho reduzido apenas se a diferença entre $\hat{\mu}_{w_0}$ e $\hat{\mu}_{w_1}$ for maior – ou igual – ao limiar ϵ_{cut} , que é estabelecido na equação 2.8,

$$\epsilon_{cut} = \sqrt{\frac{2}{m} * \sigma_w^2 * \log\left(\frac{2}{\delta'}\right)} + \frac{2}{3m} * \log\left(\frac{2}{\delta'}\right) \quad (2.8)$$

onde m é a média harmônica de n_0 e n_1 , os quais representam os tamanhos de w_0 e w_1 respectivamente. Aliás, a variância observada nos elementos de w é definida por σ_w^2 e $\delta \in [0, 1]$ refere-se ao nível de confiança (definido como 0,002 por padrão). Finalmente δ' é dado por: $\delta' = \delta / \log(n)$.

O STEPD (NISHIDA; YAMAUCHI, 2007) tem como estratégia usar duas janelas deslizantes (recente e antiga) para DMC. Ele trabalha de acordo com as seguintes suposições:

- Igualdade entres as acurácias de um classificador para w exemplos recentes e a precisão geral computada desde o início do processo de aprendizagem, desde que não haja mudança do conceito.
- A diminuição significativa na acurácias da janela recente indica o acontecimento de uma mudança de conceito.

A estatística apresentada na equação 2.9 é utilizada pelo método para realizar a comparação entre as acurácias das janelas recente e antiga. O número de predições corretas sobre os n_o exemplos da janela antiga é armazenado como r_o , excluindo da comparação os w exemplos recentes, e o r_r é o número de predições corretas sobre os w (n_r) exemplos, da janela atual, onde $\hat{p} = (r_o + r_r)/(n_o + n_r)$.

$$T(r_o, r_r, n_o, n_r) = \frac{|r_o/n_o - r_r/n_r| - 0,5(1/n_o + 1/n_r)}{\sqrt{\hat{p}(1-\hat{p})(1/n_o + 1/n_r)}} \quad (2.9)$$

A estatística de 2.9 é uma aplicação do teste de hipóteses para comparação entre proporções de duas amostras independentes. Finalmente, o critério de decisão é determinado pelo valor p encontrado por meio de tabela da distribuição normal padrão usando o resultante de $T(r_o, r_r, n_o, n_r)$. Caso o valor p seja menor do que o nível de significância adotado (α_w para *warning* e α_d para *drift*), a hipótese nula que assume a igualdade entre as acurácias das amostras independentes das janelas antiga e recente será rejeitada e o detector entrará em *warning* ou em *drift* (de acordo com o valor de α comparado) (NISHIDA; YAMAUCHI, 2007).

O ECDD (ROSS et al., 2012) foi proposto como uma adaptação do método *Exponentially Weighted Moving Average* (EWMA) (ROBERTS, 1959) com a finalidade de ocupar-se das tarefas de classificação num fluxo de dados e atuar com mais agilidade diante de possíveis mudanças de conceitos.

O método armazena as predições do classificador numa variável estabelecida como X_t e logo depois define a sua probabilidade de erro no instante de tempo t como p_t . Desse modo, é possível monitorar a detecção das prováveis mudanças de conceitos por variações em p_t .

Os níveis de *warning* e *drift* são sinalizados por ECDD por meio de inequações levando em consideração a estimativa da média dos acertos do classificador base e a sua variação de probabilidade de erro. É usado um limiar para definir a distância necessária para a detecção de uma mudança de conceito à medida em que o tempo vai passando. Por fim, $\hat{\sigma}_{Z_t}$ refere-se ao desvio padrão de Z_t e tem o seu valor definido pela equação 2.10, onde λ é um parâmetro definido para ponderar o aumento das instâncias atuais em relação às mais antigas e $\hat{\sigma}_{X_t}$ representa o desvio padrão de X_t .

$$\hat{\sigma}_{Z_t} = \sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2t})} \hat{\sigma}_{X_t} \quad (2.10)$$

O ECDD também possui o parâmetro n responsável por determinar o número mínimo de instâncias que devem ser analisadas antes que o método passe a detectar possíveis mudanças de conceitos.

O SeqDrift1 (SAKTHITHASAN; PEARS; KOH, 2013) foi projetado para ampliar sua aplicabilidade a fluxos com características diferentes, exibindo um desempenho mais

preciso e robusto em comparação com outros detectores como ADWIN. O método não tem conhecimento dos dados de distribuição subjacentes e é econômico em termos de custo computacional e consumo de memória. Também faz uso de duas sub-janelas para representar os dados recentes e antigos na distribuição do fluxo.

Para comparar as médias aritméticas das distribuições em ambas janelas, o detector utiliza o limite de *Bernstein* (BERNSTEIN, 1946), que fornece um limite bem reduzido, evitando, assim, que a estimativa da probabilidade de grandes variações para distribuições com uma variação pequena seja custosamente conservadora, como acontece com a desigualdade de *Hoeffding* (BIFET; GAVALDÀ, 2007a). Em outras palavras, a desigualdade de *Hoeffding* é considerada conservadora demais (BIFET; GAVALDÀ, 2007b) e superestima a probabilidade de grandes mudanças em distribuições com pequenas variações (SAKTHITHASAN; PEARS; KOH, 2013). Portanto, a desigualdade de *Bernstein* fornece um limite maior e é, portanto, adotada no SeqDrift1. A desigualdade de *Bernstein* está definida pela inequação 2.11.

$$Pr \left(\left| \frac{1}{n} \sum_1^n X_i - E[X] \right| > \epsilon \right) \leq 2exp \left(\frac{-n\epsilon^2}{2p\hat{\sigma}^2 + \frac{2}{3}\epsilon(c-a)} \right) \quad (2.11)$$

Observe-se que X_1, \dots, X_n são variáveis aleatórias independentes, $E[X]$ é o valor esperado ou a média da população, e $X_i \in [a, c]$, $\hat{\sigma}^2$ é a variância da amostra. Por último, ϵ é um limiar de ponto de corte obtido pelo cálculo da equação 2.12, onde b corresponde ao tamanho do bloco e $p = \ln \left(\frac{4}{\delta} \right)$.

$$\epsilon = \frac{2}{3b} \left\{ p + \sqrt{p^2 + 18\sigma_s^2 bp} \right\} \quad (2.12)$$

Em SeqDrift1, os níveis de *warning* e *drift* são estabelecidos conjuntamente com um fator de correção δ' para controlar a probabilidade de falsos positivos.

O SeqDrift2 (PEARS; SAKTHITHASAN; KOH, 2014) é uma versão estendida do SeqDrift1 apresentado na subseção anterior. O método utiliza a mesma estratégia de teste de hipóteses, mas contém uma série de melhorias importantes, incluindo a utilização de uma janela de amostragens para o gerenciamento de memória e o uso de um limite muito mais reduzido para o limiar de ponto de corte ϵ .

O limiar de ponto de corte ϵ em SeqDrift2 é denotado na equação 2.13, onde $k = n_r/n_l + n_r$ e n_l e n_r são os tamanhos dos repositórios esquerdo e direito, respectivamente.

$$\epsilon = \frac{1}{3(1-k)n_r} \left\{ p + \sqrt{p^2 + 18\sigma_s^2 n_r p} \right\} \quad (2.13)$$

O SEED (HUANG et al., 2014) acumula instâncias de dados em blocos de tamanho b . As instâncias de dados consistem em uma sequência binária de bits em que o binário 1 indica um erro de classificação incorreta e o binário 0 indica uma decisão correta de classificação feita por um classificador. São usados blocos de instâncias de dados como a

unidade básica em vez de instância a instância e assume-se que a distribuição dentro de um bloco permanece razoavelmente estável.

De acordo com Frías-Blanco et al. (2015), *Hoeffding-based Drift Detection Method* (HDDM) é uma família de métodos que se propõem a monitorar as métricas de desempenho medidas durante o processo de aprendizagem e sinalizar as mudanças quando uma variação significativa é detectada. Os métodos aplicam desigualdades de probabilidade que assumem apenas variáveis aleatórias independentes, uni-variáveis e delimitadas para obter garantias teóricas para as detecções.

O **A**Test é implementado mediante modificações realizadas a um corolário proposto por Hoeffding (1963). Este pode ser aplicado à detecção de mudanças significativas nas médias móveis dos valores de um fluxo de dados, surgindo, assim, um teste estatístico bicaudal onde são detectadas as mudanças na média da população e, por conseguinte, permite monitorar a diferença entre as médias abordadas do teste utilizando uma sequência de variáveis aleatórias $X_1, \dots, X_n, Y_1, \dots, Y_m$. O método tem um melhor comportamento na detecção de mudanças de conceito abruptas.

O HDDMW é o segundo método da família do HDDM (FRÍAS-BLANCO et al., 2015). O **W**Test é baseado em um teste estatístico mais geral que usa médias móveis ponderadas, embora igualmente eficiente e simples. Neste caso, os valores recentes no fluxo de dados têm mais peso do que os mais velhos, assumindo que eles têm maior probabilidade de ocorrência. O teste McDiarmid (1989) é uma generalização do conceito de *Hoeffding bound* para variáveis aleatórias dependentes. Finalmente, este método é mais recomendável para a utilização das execuções com as mudanças de conceito graduais.

O FHDDM (PESARANGHADER; VIKTOR, 2016) foi proposto com a finalidade de minorar a taxa de FP e FN ao detectar as alterações na distribuições dos dados. O algoritmo desliza uma janela com um tamanho de n (200 valor padrão) nos resultados da classificação. Posteriormente, insere um 1 na janela se o resultado da previsão for verdadeiro, no caso contrário insere 0. À medida que as entradas são processadas, calcula a probabilidade de observar 1s (p_t^1) na janela deslizante no tempo t , e também mantém a probabilidade máxima de ocorrência de 1s, ou seja, p_{max}^1 . Na equação 2.14 mostra-se que se o valor de p_1 no tempo t é maior do que o valor de p_{max}^1 , então o valor de p_{max}^1 será atualizado.

$$if p_{max}^1 < p_t^1 \Rightarrow p_t^1 \rightarrow p_{max}^1 \quad (2.14)$$

Assim, os autores demonstram, sob a base do modelo de aprendizado *Probably Approximately Correct* (PAC) (MITCHELL, 1997), que a possibilidade de enfrentar uma mudança de conceito aumenta se p_{max}^1 não muda e p_t^1 diminui ao longo do tempo. Eventualmente, como apresenta-se na equação 2.15, uma diferença significativa entre p_{max}^1 e p_t^1 indica a ocorrência da mudança de conceito no fluxo de dados, onde ε_d é

determinado usando a probabilidade do erro δ (padrão 10^{-7}) fornecida pelo conceito de *Hoeffding bound* (HOEFFDING, 1963; MARON; MOORE, 1993).

$$\Delta p = p_{max}^1 - p_t^1 \geq \varepsilon_d \Rightarrow \text{Drift} := \text{True} \quad (2.15)$$

O FHDDMS (PESARANGHADER; VIKTOR; PAQUET, 2018) estende FHDDM mantendo janelas de tamanhos diferentes, uma janela deslizante curta e longa. A lógica por trás dessa abordagem é reduzir o atraso na detecção e a taxa de falsos negativos. Intuitivamente, uma janela curta deve detectar mudanças de conceitos abruptas mais rapidamente, enquanto uma janela longa deve detectar mudanças de conceitos graduais com uma taxa de falsos negativos mais baixa.

O RDDM (BARROS et al., 2017) periodicamente encurta o número de instâncias de conceitos estáveis muito longos para lidar com um problema conhecido de perda de desempenho do DDM. Com este propósito, descarta instâncias antigas de conceitos muito longos com o objetivo de detectar desvios mais cedo, melhorando a precisão de suas detecções e, principalmente, a precisão final. Além disso, usando sua configuração padrão recomendada, o RDDM apresenta um desempenho especialmente forte em conjuntos de dados com mudanças de conceito graduais e quando os tamanhos dos conceitos possuem milhares de instâncias.

O RDDM tem foco no problema de perda da eficiência que ocorre no DDM devido a diminuição da sensibilidade quando há conceitos longos. Periodicamente, por meio de janela deslizante, são descartados valores antigos de acertos e erros da classificação. Após isso, são recalculadas as estatísticas responsáveis por alertas de *warning* e *drifts*, além disso, também força a ocorrência de *drift* (falso *drift* – *drift* interno) quando o conceito alcança limites para serem considerados como longos. Por fim, na maioria das vezes o RDDM melhora a precisão da classificação quando comparado os resultados de desempenho com o DDM, especialmente em conjuntos de dados com mudanças graduais, e, faz isso detectando as mudanças mais cedo, apesar do aumento moderado do número de FP.

O WSTD (BARROS; HIDALGO; CABRAL, 2018) é um eficiente detector de mudanças de conceito que surgiu com o objetivo de identificar menos FP do que STEPDP e que também fosse estatisticamente mais preciso. Para isso, a implementação do método utiliza o teste estatístico da soma dos ranks de *Wilcoxon*, provêm de populações com a mesma distribuição nos dados (LARSON; FARBER, 2010).

Este detector funciona de forma semelhante a STEPDP, monitorando as predições do classificador e para isso ele utiliza duas janelas deslizantes (recente e antiga). Uma das principais diferenças entre STEPDP e WSTD, está relacionada com o tamanho da janela deslizante mais antiga, para o qual no WSTD o valor é fixo ao invés de todos os outros exemplos, como é usado no STEPDP.

Na equação 2.16, é apresentada a estatística do método utilizada para comparar as distribuições entre as duas janelas, onde R é a soma dos postos para a menor amostra. Neste sentido, o cálculo dos postos foi simplificado matematicamente no método, usando a fórmula para calcular a soma dos elementos das séries aritméticas (AS) - progressões aritméticas finitas, tornando desnecessário o uso de uma ordenação explícita como acontece no teste estatístico original, devido ao fato de que as observações da classificação são binárias (0 ou 1).

$$z = \frac{R - \frac{n_r(n_o+n_r+1)}{2}}{\sqrt{\frac{n_o n_r (n_o+n_r+1)}{12}}} \quad (2.16)$$

Por último, o critério de decisão determinado pelo valor p para rejeitar ou não a hipótese nula e, por conseguinte, estabelecer os estados de *warning* ou *drift* do detector, é realizado da mesma maneira como procede o STEPD.

O FTDD (CABRAL; BARROS., 2018) é um dos três detectores de mudança de conceito baseados em uma implementação do teste Exato de *Fisher*. Baseia-se no STEPD (NISHIDA; YAMAUCHI, 2007) e na deficiência do seu teste estatístico de proporções iguais em situações em que as amostras de dados são pequenas ou desequilibradas. O FTDD detecta alterações na distribuição de dados usando exclusivamente o teste exato de *Fisher*. Seu cálculo do valor de p é o mais simples dos três métodos. Vale ressaltar que a ordem das operações no cálculo do valor de p do FTDD foi alterada, na tentativa de evitar valores extremos nos resultados intermediários, como medida de precaução para tornar os resultados o mais precisos possível.

O FPDD visa usar o teste exato de *Fisher* nas situações em que o número de erros ou o número de previsões corretas, em qualquer uma das duas janelas, é pequeno. Caso contrário, ele usa o teste de proporções iguais, assim como o STEPD. Para decidir se o número de erros e acertos é pequeno ou não, foi necessário escolher o valor mínimo que aplicaria o teste de proporções iguais.

Da mesma forma que o FPDD, o FSDD também aplica o teste exato de *Fisher* quando o número de erros ou previsões corretas em qualquer uma das duas janelas é menor que cinco. No entanto, nos demais cenários, o FSDD adota o teste estatístico do qui-quadrado para homogeneidade de proporções, em vez do teste de proporções iguais usado no STEPD e no FPDD.

2.2.3 Detectores de Múltipla Análise

O comitê de detectores *e-Detector* (DU et al., 2014) foi desenvolvido para localizar mudanças de conceitos abruptas e graduais. Inicialmente se aplicam técnicas para juntar os detectores de acordo com a sua diversidade, e selecionar o número de detectores para fazer parte do comitê (ZHOU; WU; TANG, 2002). O *e-Detector* consiste das fases de aprendizagem online e a detecção de mudança de conceito.

Para medir a diversidade entre dois detectores, os autores definiram dois tipos de indicadores mudança: homogêneo e heterogêneo. Se dois detectores possuem estatísticas equivalentes em encontrar mudanças de conceitos, se diz que são homogêneos e do contrário são heterogêneos.

Sendo assim, a diversidade é obtida observando a estatística que se refere ao indicador de mudança de conceito do detector. Por exemplo, a estatística do DDM é obtida pela taxa de erro, já no STEPDP ela se obtém a partir da precisão, diferentemente, o EDDM calcula em um intervalo entre erros de classificação. Vale ressaltar que, a taxa de erro e erros de classificação são equivalentes, enquanto que a medida precisão leva em consideração também as classificações corretas. Logo, os métodos DDM e STEPDP são homogêneos, enquanto os métodos DDM e EDDM são heterogêneos, consequentemente EDDM e STEPDP também são considerados heterogêneos.

O ensemble baseia-se na regra “O primeiro que encontra é o primeiro que reporta” para assumir os três possíveis estados:

- *Controlado*: todos os detectores base do comitê reconhecem que o fluxo é estável;
- *Alerta*: se ao menos um dos detectores notificar sinal de alerta;
- *Mudança*: se precisa apenas que um dos detectores notifique a mudança de conceito.

Em Wozniak et al. (2016), são apresentadas três regras para combinar detectores heterogêneos. Como se mostra na equação 2.17, eles assumem que se tem um conjunto de n detectores de mudanças.

$$\mathbb{D} = \{D_1, D_2, \dots, D_n\} \quad (2.17)$$

Para garantir a heterogeneidade, cada um dos detectores escolhidos tem a possibilidade de emitir um de três sinais possíveis, sendo:

$$D_i = \begin{cases} 0 & \text{se a mudança não é detectada} \\ 1 & \text{se o nível de warning (*alerta*) é atingido} \\ 2 & \text{se a mudança é detectada} \end{cases} \quad (2.18)$$

As regras de decisão do comitê apresentadas somente levam em consideração a notificação do estado de mudança de conceito dos detectores que conformam o comitê. Essas regras podem ser enumeradas da seguinte maneira:

- ALO: O comitê assume a ocorrência da mudança de conceito se ao menos um dos detectores informa a ocorrência da mudança;

- ALHD: A mudança de conceito é declarada se a metade dos detectores do comitê a informam;
- AD: Todos os detectores do comitê notificam a ocorrência da mudança.

Para cada uma das regras anteriores foi construído um comitê de detectores de mudanças de conceito. Segundo os autores Wozniak et al. (2016), os comitês apresentados nesta parte de experimentação, para provar o funcionamento das regras, não têm ótimo rendimento já que a escolha dos detectores que conformam os comitês foi realizada sem uma profunda análise, pois seu objetivo somente consistiu em propor novas metodologias para a criação de comitês (ensembles).

2.3 GERADORES E BASES DE DADOS REAIS

Em informática, o termo gerador é utilizado para denotar programa que gera outro programa, seguindo as especificações recebidas. No caso dos geradores de dados, diz respeito de programa que gera dados, também seguindo as especificações. Assim sendo, o conteúdo gerado por uma especificação pode respeitar um modelo de dados que descreve algo por meio de simulação.

Os termos “base de dados”, “coleção de dados”, “repositório de dados” e “conjunto de dados” são usados como sinônimos em AM para representar um repositório de informação. Eventualmente, esses repositórios são alimentados por fatos armazenados por sistemas de informação, e/ou eventos baseados em sensores que registram algo do mundo real. Diz-se base de dados real ao repositório de informação populado por conteúdo produzido sob certa casualidade do mundo real, ou seja, não por simulação.

Os geradores de dados *Agrawal*, *Light Emitting Diode* (LED), *Random Radial Basic Function* (RRBF), *Sine*, *WaveForm* e *Mixed* usados nesta tese estão disponíveis no MOA. Eles foram utilizados em vários trabalhos na literatura, como Santos, Barros e Gonçalves Jr. (2019), Barros e Santos (2019), Hidalgo, Maciel e Barros (2019) bem como as bases de dados reais *Airline*, *Connect-4*, *SpamData*, *Outdoor*, *Rialto* e *Wine-Quality-Red*.

2.3.1 Geradores de Dados

Os geradores são modelos de simulações usados para criar situações mediante criação de dados artificialmente. Nesta tese, eles são utilizados nas avaliações de desempenho para verificar o comportamento e analisar os pontos fortes e fracos dos algoritmos. Nesse sentido, ao analisá-los, é possível adotar estratégias que possam ser utilizadas para melhor lidar com as situações. A Tabela 1 apresenta um breve resumo sobre os geradores de dados.

O gerador *Agrawal* (AGRAWAL; IMIELINSKI; SWAMI, 1993; SANTOS et al., 2014) armazena informações de pessoas com o objetivo de receber um empréstimo, e essas pessoas são classificadas como grupo A ou B. É uma classificação binária do perfil da

Tabela 1 – Resumo descritivo dos geradores. No gerador RRBF a quantidade de conceitos é representada por n que pode variar de 1 à maior que 1 milhão.

| # | Gerador | Atributo | Classe | Conceito |
|---|----------|----------|--------|----------|
| 1 | Agrawall | 9 | 2 | 10 |
| 2 | LED | 24 | 10 | 7 |
| 3 | Mixed | 4 | 2 | 2 |
| 4 | RBF | 10 | 2 | n |
| 5 | Sine | 2 | 2 | 4 |
| 6 | Waveform | 21 | 3 | 21 |

Fonte: O AUTOR (2020).

pessoa se ela deve ou não receber empréstimo. Os atributos são: salário, comissão, idade, nível de instrução, código postal, etc. Possui dez funções, cada uma com diferentes formas de avaliação. Além disso, por padrão, é adicionado ruído aos dados gerados. Mudanças de conceito são simuladas alternando as funções de avaliação.

O gerador LED (light-emitting diode) (GONÇALVES JR.; BARROS, 2013b; FRÍAS-BLANCO et al., 2015) representa o problema de prever o dígito de um diodo emissor de luz de sete segmentos, também conhecido como *display* LED. Ele possui 24 atributos categóricos (17 deles são irrelevantes) e uma classe categórica. Cada atributo tem uma probabilidade de 10% de ser invertido (nível padrão de ruído). Mudanças de conceito são simulados alterando a posição dos sete atributos relevantes.

O gerador *Mixed* (GAMA et al., 2014, p.6) cria dados compostos por 4 atributos, dos quais 2 são booleanos (v e w) e 2 são numéricos (x e y). As instâncias são classificadas em positivas ou negativas, dependendo dos valores dos atributos. Para que uma instância seja positiva, a regra diz que pelo menos duas das seguintes condições devem ser verdadeiras: os valores de v e w devem ser verdadeiros e $y < 0,5 + 0,3 \sin(3\pi x)$. As mudanças de conceitos são simuladas invertendo as classes. Para que uma instância seja considerada positiva, pelo menos duas das seguintes condições devem ser verdadeiras: os valores de v e w deverão ser falsos e $y \geq 0,5 + 0,3 \sin(3\pi x)$.

O gerador RRBF (BIFET et al., 2009; SANTOS; BARROS; GONÇALVES JR., 2015) usa n centroides com seus centros, classes e pesos definidos aleatoriamente e uma distribuição gaussiana para determinar os valores de a atributos. Mudanças de conceito podem ser simuladas mudando as posições dos centroides.

O gerador da base de dados *Sine* (GAMA et al., 2004; BARROS; SANTOS, 2019) usa dois atributos numéricos (x, y) , duas classes (“negativo” e “positivo”, Sine1 e Sine2, respectivamente) e quatro funções para gerar os conceitos. Em Sine1, cada instância é positiva se o ponto (x, y) estiver abaixo da curva $y = \sin(x)$, enquanto Sine2 usa a função $y = 0,5 + 0,3 \times \sin(3\pi x)$. Mudanças de conceito são simuladas alternando entre funções de conceito.

O *WaveForm* (BIFET; HOLMES; PFAHRINGER, 2010; MACIEL; SANTOS; BARROS, 2015) é composto por três classes e 40 atributos numéricos, onde os últimos 19 são irrelevantes para o conceito (são ruídos). Neste gerador cada classe é definida a partir da mistura de duas ondas. Existem três ondas, sendo possível três combinações entre elas. Mudanças de conceito são simuladas alternando as posições dos atributos que representam o conceito.

2.3.2 Base de Dados Reais

Em relação às bases de dados reais, normalmente não há uma indicação exata em relação às mudanças de conceito. Conseqüentemente, não é possível precisar quando a mudança começa a ocorrer; qual a velocidade da mudança; ou se existe alguma mudança de conceito. Portanto, não é possível realizar uma análise detalhada do comportamento dos algoritmos na presença de mudança de conceito usando apenas bases reais sem essas garantias em relação às mudanças de conceito. A Tabela 2 apresenta um breve resumo sobre as bases de dados reais.

Tabela 2 – Resumo descritivo das bases de dados reais.

| # | Base de dados | Atributo | Classe | Total de Exemplos |
|---|--------------------|----------|--------|-------------------|
| 1 | Airlines | 7 | 2 | 539.383 |
| 2 | Connect-4 | 42 | 3 | 67.557 |
| 4 | Outdoor | 10 | 40 | 4.000 |
| 5 | Rialto | 27 | 10 | 82.250 |
| 3 | SpamData | 500 | 2 | 9.324 |
| 6 | Wine Quality Red | 11 | 9 | 1.599 |
| 7 | Wine Quality White | 11 | 9 | 4.898 |

Fonte: O AUTOR (2020).

A *Airlines* (IKONOMOVSKA; GAMA; DŽEROSKI, 2011; BARROS; SANTOS, 2019) é um conjunto de dados binários com 539.383 instâncias, inspirado no conjunto de dados de regressão proposto por (IKONOMOVSKA, 2008; IKONOMOVSKA; GAMA; DŽEROSKI, 2011). O conjunto de dados contém detalhes de chegada e partida de voos comerciais nos Estados Unidos da América. São 7 atributos: nome da empresa (com 18 possibilidades), número do voo, aeroportos de origem e destino (com 293 possibilidades), dia da semana, horário do voo e horário de chegada. O objetivo é prever se os voos estão atrasados ou não, dadas as informações da partida programada.

Connect-4 (TROMP, 1995) contém todas as posições legais de 8 camadas no jogo de *connect-4* em que nenhum dos jogadores ganhou ainda e no qual o próximo lance não é forçado. A base é composta por 42 atributos e 67.557 instâncias. A classe de resultado é o valor teórico do jogo para o primeiro jogador: ganhar, perder ou empatar.

Outdoor (LOSING; HAMMER; WERSING, 2016; BARROS; HIDALGO; CABRAL, 2018) é uma base obtida a partir de imagens gravadas por um telefone celular em um ambiente de jardim. A tarefa é classificar 40 objetos diferentes; cada um deles foi observado cinco vezes em sol e cinco vezes em condições nubladas. As abordagens foram realizadas em diferentes direções. Cada abordagem representa 10 imagens em ordem temporal dentro do conjunto de dados, para um total de 4.000 instâncias.

Rialto (LOSING; HAMMER; WERSING, 2016) é um conjunto de dados contendo 82.250 instâncias e dez classes. Ela representa dois edifícios coloridos perto da famosa Ponte *Rialto* em Veneza. As cores foram codificadas em um histograma normalizado de 27 dimensões RGB (atributos). As imagens foram obtidas a partir do espaço de tempo em fotos da *webcam* de vídeo fixo. As gravações cobrem 20 dias consecutivos durante os meses de maio e junho de 2016, sendo que as mudanças contínuas no clima e nas condições de iluminação afetam o desempenho.

SpamData (KATAKIS; TSOUMAKAS; VLAHAVAS, 2010; KATAKIS G. TSOUMAKAS, 2008) é um conjunto de dados de alta dimensionalidade baseado nas mensagens de e-mail da coleção original de *Spam Assassin* (FOUNDATION, 2006). A abordagem foi utilizada para representar e-mails e possui 9.324 instâncias e 500 atributos (palavras derivadas após a seleção de recursos). Os autores afirmam que as características dessas mensagens de spam mudam gradualmente com o tempo, ou seja, esse conjunto de dados contém mudança de conceito gradual.

Os conjuntos de dados *Wine-Quality-Red* e *Wine-Quality-White* são apresentados por Cortez et al. (2009), baseados em dados de vinhos (vermelho e branco). Devido a questões de privacidade e logística, eles contêm apenas as variáveis físico-químicas (entradas) e sensoriais (saída). Os dados sobre tipos de uvas, marcas de vinho, preços e algumas outras informações não estão incluídos.

As bases de dados *Wine-Quality-Red* e *Wine-Quality-White* possuem 1.599 e 4.898 instâncias, respectivamente. Ambas têm 11 atributos mais a saída (classe) que é baseada em dados sensoriais (mediana de pelo menos 3 avaliações feitas por especialistas em vinho). Cada especialista classificou a qualidade do vinho entre 0 (muito ruim) e 10 (excelente), entretanto, não existe classe zero e dez na base de dados. O objetivo é modelar a qualidade do vinho com base em testes físico-químicos.

2.4 APRENDIZADO ADAPTATIVO COM DETECTOR EMBUTIDO

Aplicar algoritmos de AM tradicionais em ambientes com variáveis não estacionárias pode ser visto como triste ou animador. É triste porque faz com que a pesquisa feita no passado fique obsoleta, já que os algoritmos projetados não são adequados para resolver os novos desafios gerados pelo ambiente de fluxo de dados. Além disso, AM em fluxo de dados requer um conjunto diferente de habilidades computacionais (como apresentado na sub-

seção anterior) daquelas usadas no AM tradicional, uma destas habilidades é identificar a ocorrência de mudanças de conceito e tomar alguma ação corretiva. Por outro lado, a DMC é empolgante porque traz uma infinidade de novos desafios, alguns já conhecidos e alguns ainda a serem descobertos. Como esses novos desafios precisarão ser resolvidos, o campo de DMC está trazendo um novo dinamismo aos campos de Mineração de Dados (do inglês, *Data Mining*) (MD) e AM.

No entanto, o que realmente é DMC? -Basicamente, como definição geral, uma MC em AM refere-se a mudanças no conjunto de regras que definem as associações dos atributos de entrada e saída para um conceito. Conseqüentemente, em relação aos impactos no conjunto de dados, a MC implicará uma nova maneira de confeccionar os padrões de exemplos que serão atribuídos às classes do novo conceito. Do ponto de vista do aprendizado, um dos efeitos da MC no modelo de classificação está na perda da capacidade de associar entradas e saídas corretamente. Portanto, **DMC** refere-se à tarefa de *identificar a presença de um novo conceito no fluxo de dados*.

Definição 3 – Conjunto de Dados – refere-se a uma coleção de dados. Na estatística são tratados de duas maneiras, conjunto população ou conjunto amostra.

Definição 4 – População – é uma coleção de todos os resultados, respostas, medições ou contagens que são de interesse.

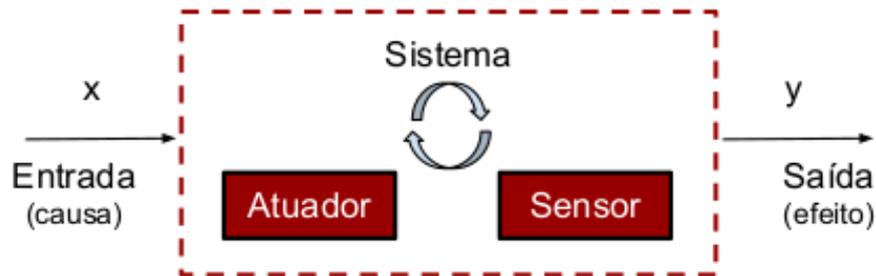
Definição 5 – Amostra – é um subgrupo de uma população.

O termo *deteção* em DMC refere-se ao fato de o algoritmo de DMC trabalhar embutido no classificador e sua tarefa ser semelhante à de um sensor. Em geral os sensores são dispositivos usados para detectar/captar estímulos externos, como por exemplo fenômenos físicos tais como frio, temperatura, etc. Neste sentido, os detectores por meio de informações do ambiente, atuam detectando a presença de MC.

A Figura 3 simboliza um sistema baseado em sensor e atuador, algo equivalente a um sistema baseado em AADE. Em geral os atuadores fazem o inverso aos detectores, isto é, ao invés de transformarem uma grandeza física em algo, eles transformam algo em energia física. Nessa abstração de sistema de atuador e sensor, o atuador ilustra o modelo, pois ele interage com o meio (construindo/atualizando o modelo). Neste sentido, o sensor (detector) interage com o sistema para que ele adapte o atuador (indutor) à condição atual do fluxo de dados.

O Quadro 2, ilustra uma separação baseada em uma pilha de camadas do modelo de referência de AADE (GAMA et al., 2004). Cada camada opera de modo independente uma da outra. A separação em camadas contribui para um melhor entendimento das partes do problema e suas respectivas responsabilidades. Essas camadas operam entregando valor da camada inferior para a camada superior. Neste sentido, a *Camada de Dados* é responsável

Figura 3 – Agente. Abstração de sistema básico de entrada e saída baseado em atuador e sensor. O objetivo é fazer um comparativo entre os papéis do sistema e algoritmo de adaptação, sensor e algoritmo de DMC, assim como o atuador em relação ao algoritmo de AM. As entradas x podem ser comparadas aos dados em AADE e saídas y como os resultados das classificações.



Fonte: O AUTOR (2020).

por entregar uma instância à *Camada Classificador*, que, por sua vez, informa à *Camada Detector* se obteve sucesso ou insucesso na previsão, e finalmente a *Camada Adaptador* recebe da *Camada Detector* um diagnóstico em relação as respostas do classificador. O Quadro 1 apresenta os diagnósticos e nomenclaturas utilizadas pela comunidade de MC, segundo o padrão de variação ao longo do tempo, tratando-se dos acertos e erros do classificador. Na *Camada Adaptador*, ocorre a estratégia de adaptação do modelo ao fluxo de dados.

Quadro 1 – Categorias dos estados de detecção de mudança de conceito.

| Estado | Detalhe |
|-------------------------|---|
| Stable (estável) | Converge para flutuações regulares em torno de um valor fixo. |
| Warning (alerta) | Apresenta oscilações periódicas. |
| Drift (mudança) | Estabelece-se por variações irregulares. |

Fonte: O AUTOR (2020).

Quadro 2 – Resumo das camadas do Modelo de referência de AADE.

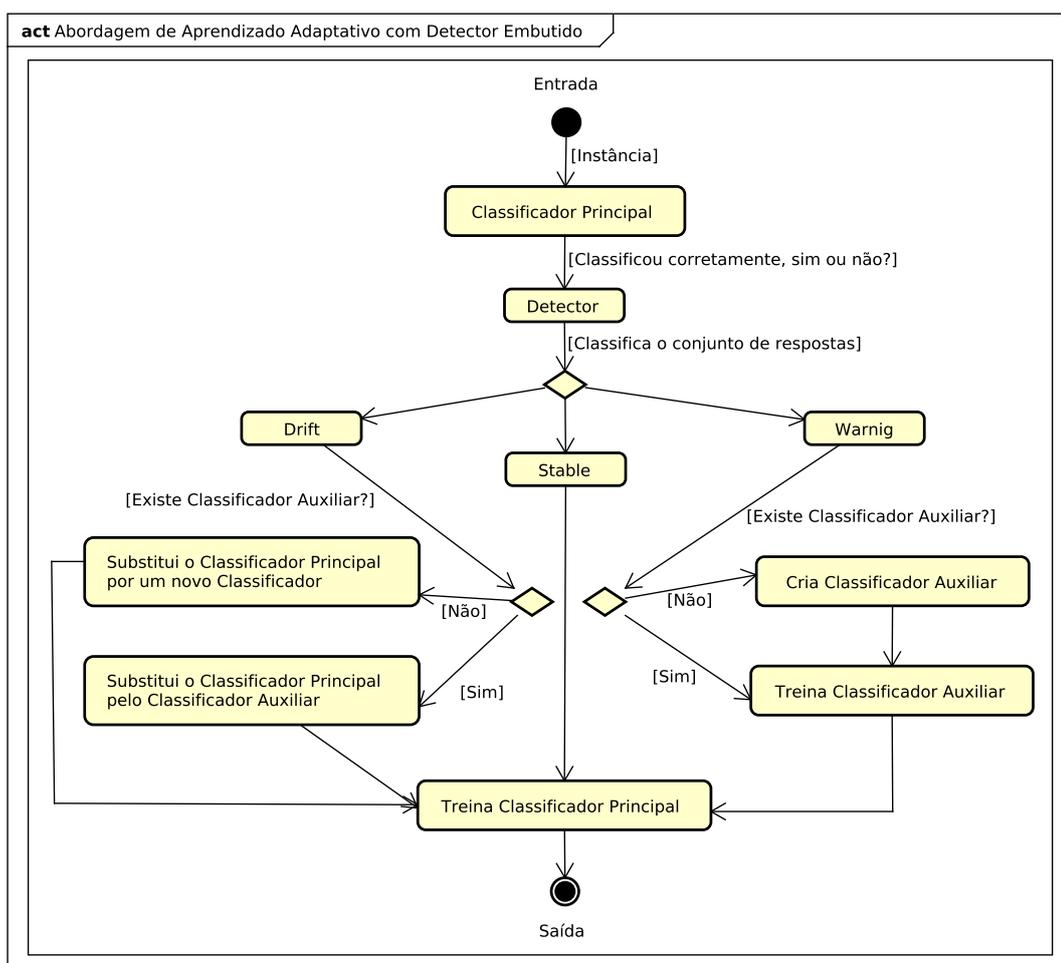
| Camada | Nome | Detalhe |
|--------|---------------|---|
| 4 | Adaptador | Fazer adaptações ao modelo de aprendizado. |
| 3 | Detector | Algoritmo de detecção de mudança de conceito. |
| 2 | Classificador | Algoritmo de aprendizado de máquina. |
| 1 | Dados | Fluxo de dados a ser classificado. |

Fonte: O AUTOR (2020).

Por fim, a Figura 4, apresenta em forma de diagrama de atividade o funcionamento da abordagem AADE. A tarefa tem início na instância a ser classificada, seguidamente,

o detector é acionado. Dada a situação diagnosticada pelo detector, existem apenas três alternativas. Por fim, o resultado é a construção/atualização do modelo principal. Na próxima subseção, são explorados exemplos de cenários reais que podem ser bem-sucedidos com aplicação de técnicas de DMC.

Figura 4 – Abordagem de Aprendizado Adaptativa com Detector Embutido.



Fonte: O AUTOR (2020).

2.5 CONSIDERAÇÕES FINAIS

Aplicações de MC geralmente visam automatizar alguma tarefa de tempo real, por exemplo, a operação de semáforos de uma cidade, que pode mudar de acordo com o horário do dia. Os dados fornecidos para classificação podem ser provenientes de sensores, redes sociais, aplicações para avaliação de crédito, operações financeiras do mercado de ações, telecomunicações, entre outras, e o objetivo na maior parte das vezes está relacionado a descrever algo. Sua aplicabilidade pode ser resumida em fornecer alertas ou operar de forma autônoma, identificando e corrigindo possíveis problemas.

Mais exemplos de aplicações de MC podem ser encontrados em Japkowicz (2016, p.91-110).

A ideia de DMC apresentada tem seu sentido prático. O principal assunto desta seção foi tratar do aprendizado em fluxo de dados com mudança de conceito. Por fim, foi apresentada e explicada a abordagem de Aprendizado Adaptativo com Detector Embutido, usada em alguns dos trabalhos de detectores disponíveis na literatura.

3 ASPECTOS METODOLÓGICOS

Esta pesquisa de doutorado foi desenhada com base no método científico. Esse método é composto por um conjunto de procedimentos que permitem operacionalizar o trabalho para que seja possível ter acesso às relações de causas e efeitos, ou seja, é por meio da observação dos fatos que é possível questionar os ocorridos e a partir desses achados problematizá-los por meio de pontos de vista. Assim sendo, o método científico contribuiu não apenas com as regras e procedimentos para a realização da pesquisa, mas também teve sua contribuição para apoiar o saber científico.

Dessa forma, a problemática da pesquisa parte da curiosidade em conhecer o estado da arte em detecção de mudança de conceito por meio dos trabalhos disponibilizados na literatura, ferramentas e através da formulação de hipóteses de melhorias. Nota-se que para chegar a tal objetivo foi realizada uma extensa pesquisa bibliográfica, que foi útil tanto para embasar os conceitos da pesquisa como para fundamentar os experimentos e análises de dados. A análise, organização, interpretação e comparação desses dados foram feitas, principalmente, mediante técnicas estatísticas, que contribuíram para identificação das diferenças entre amostras, classificação ordenada de desempenho e para uma melhor compreensão dos métodos, identificando pontos fortes e fracos.

Para investigar os pontos fortes e fracos das abordagens de DMC, utilizou-se de uma pesquisa com objetivo exploratório e descritivo do tema. A parte empírica foi embasada com pesquisa teórica, fazendo uso de uma pesquisa com procedimentos experimentais. Assim sendo, foi incluída a realização de muitos procedimentos experimentais e posterior análise de dados, conforme técnicas estatísticas de descrição e análise de dados que permitem avaliar afirmações sobre as médias de populações.

3.1 ANÁLISE DE DADOS

A análise de dados reúne várias etapas para extrair significado dos resultados obtidos. No que se refere a essas diferentes etapas, optou-se por organizar em três fases: 1) pré-análise; 2) exploração dos resultados; e 3) tratamento dos resultados, inferência e interpretação.

3.1.1 Pré-Análise

Esta subseção é responsável pela organização do material a ser analisado de modo sistemático. Em resumo, trata-se da organização dos dados por meio de quatro etapas: (a) escolha dos dados, que consiste na demarcação do que será analisado; (b) extração dos resultados, que diz respeito à coleta de dados e, à organização dos dados em documentos; (c) formulação das hipóteses e dos objetivos; (d) elaboração de indicadores.

Para experimentos com dados artificiais, adotaram-se 40 repetições. De acordo com o teorema central do limite, dado um conjunto de amostras maior ou igual a 30 (suficientemente grande), é possível realizar estimação de aproximação em relação à distribuição normal. Entretanto, ter um número maior de amostras aumenta as garantias teóricas em relação às aproximações dos resultados e permite maximizar o poder dos testes estatísticos. Logo, adota-se um número 1/3 maior (igual a 40 repetições) que o mínimo (30) como padrão para coleta de amostras.

A escolha dos dados consiste na seleção dos trabalhos relacionados de detectores a serem utilizados nas configurações gerais dos experimentos adotados nesta tese. Para realização da avaliação empírica, as abordagens apresentadas na seção de trabalhos relacionados foram adotados. Em cada experimento, que consiste em um conjunto de parametrizações ou simplesmente *script* MOA, há uma combinação de seis elementos básicos: 1. método de aprendizado; 2. método detector; 3. gerador de dados; 4. velocidade da mudança; 5. tamanho da base de dados; e 6. pontos de mudanças.

Em relação ao elemento 1 (método de aprendizado), foram adotados dois:

- *Hoeffding Tree* (HT);
- *Naive Bayes* (NB).

Tratando-se do item 2 (método detectores), foram adotados dezesseis de única análise:

- DDM (GAMA et al., 2004);
- EDDM (BAENA-GARCIA et al., 2006);
- STEPDP (NISHIDA; YAMAUCHI, 2007);
- ECDD (ROSS et al., 2012);
- SeqDrift1 (SAKTHITHASAN; PEARS; KOH, 2013);
- SeqDrift2 (PEARS; SAKTHITHASAN; KOH, 2014);
- SEED (HUANG et al., 2014);
- HDDMA, HDDMW (FRÍAS-BLANCO et al., 2015);
- FHDDM (PESARANGHADER; VIKTOR, 2016);
- RDDM (BARROS et al., 2017);
- WSTD (BARROS; HIDALGO; CABRAL, 2018);
- FTDD, FSDD, FPDD (CABRAL; BARROS., 2018);
- FHDDMS (PESARANGHADER; VIKTOR; PAQUET, 2018).

Ainda tratando-se do item 2, foram adotados quatro de múltipla análise:

- *e-Detector* (DU et al., 2014);
- ALO, ALHD, AD (WOZNIAK et al., 2016).

Quanto ao item 3 (geradores de dados), foram adotados seis:

- *Agrawal* (AGRAWAL; IMIELINSKI; SWAMI, 1993);
- *Light Emitting Diode* (LED) (GONÇALVES JR.; BARROS, 2013b);
- *Mixed* (GAMA et al., 2014);
- *Random Radial Basic Function* (RRBF) (BIFET et al., 2009);

- *Sine* (GAMA et al., 2004);
- *Waveform* (BIFET; HOLMES; PFAHRINGER, 2010).

Quanto ao item 4 (velocidade da mudança), foram adotadas duas velocidades que simulam mudanças abruptas e graduais:

- 0 (zero);
- 500.

Os tamanhos 20k, 130k, 890k foram definidos com base na sequência de *Fibonacci*, começando por 0, na qual, cada termo subsequente corresponde à soma dos dois anteriores, ou seja, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55 e 89 definem os tamanhos totais das bases de dados. A escolha dos tamanhos com base nessa sequência foi realizada na expectativa de que exista uma proporção lógica entre aumento dos tamanho dos conceitos e eficiência dos algoritmos. Os tamanhos em si foram selecionados saltando 4 elementos da sequência, começando do zero e cada elemento selecionado é multiplicado por 10k. A sequência fica: $2 \times 10k = 20k$; $13 \times 10k = 130k$; e $89 \times 10k = 890k$.

Quanto ao item 6 (pontos de mudanças), foram adotadas quatro mudanças para cada base de dados, independentemente do tamanho, totalizando cinco conceitos por base de dados gerada. A seguir, são apresentados os tamanhos de base de dados e seus respectivos pontos de mudanças.

- 20k com mudança nos pontos 4k, 8k, 12k e 16k;
- 130k com mudança nos pontos 26k, 52k, 78k e 104k;
- 890k com mudança nos pontos 178k, 356k, 534k e 712k.

Sendo assim, para cada combinação de método de aprendizado e detector, há 6 (geradores) \times 2 (velocidades de mudança) \times 3 (tamanhos de base de dados) = 36 parametrizações. Levando em consideração apenas os detectores de única análise (16), são parametrizados $16 \times 36 = 576$ *scripts* MOA para cada método de aprendizado. Como são utilizados dois, há um total de $576 \times 2 = 1152$ possibilidades. Vale ressaltar que, para cada uma das 1152 possibilidades, há 40 repetições, logo, $1152 \times 40 = 46080$ execuções.

Por meio da avaliação empírica, procura-se avaliar um conjunto de características divididas em três componentes básicos que permitem determinar o grau de eficiência e as possibilidades de operação de determinado método detector. Os componentes básicos são divididos em duas categorias que avaliam a eficiência do (1) método de aprendizado e (2) método detector. Essa avaliação é um recurso valioso para comprovar a relação de desempenho entre os trabalhos relacionados, já que permite mensurar o desempenho dos métodos nos cenários utilizados.

Os componentes básicos apresentados fazem parte da política de avaliação de desempenho que evidencia o compromisso deste trabalho com os demais. Assim sendo, para avaliar o desempenho dos métodos de aprendizado e detector, são utilizadas

métricas baseadas em critérios combinados. O método de aprendizado é avaliado por meio da métrica de acurácia da classificação. O método detector é avaliado por meio de duas métricas: MTR e MCC (explicadas no Apêndice A). Vale ressaltar que, foram adotadas duas medidas de desempenho para avaliar o método detector devido ao tema do trabalho ser baseado em detecção de mudança de conceito. Entretanto, entende-se que um bom método detector deve obter um compromisso justo entre corretas detecções e acurácia da classificação.

É importante destacar que, durante a transição de conceitos (explicado na seção 2), não há como oferecer garantias teóricas do aprendizado (CORTES; VAPNIK, 1995). Conseqüentemente, detectar corretamente uma MC nessa condição maximiza as dificuldades tanto para rastrear o ponto exato parametrizado da mudança quanto para avaliar a corretude da detecção. A métrica MTR é utilizada para calcular a eficiência do detector em transições abruptas e graduais, apesar do questionamento teórico em relação ao correto ponto da MC. A métrica MCC é utilizada para avaliar a fidedignidade das detecções, considerando apenas transições abruptas. Sendo assim, os falsos alarmes são calculados levando em consideração a métrica MCC, entretanto, a métrica MTR é utilizada como complemento.

Tratando-se da extração dos resultados e organização, os dados foram organizados em planilhas de texto. Foram catalogadas todas as medidas que são utilizadas nas avaliações, apresentados no tópico 3.2.4. Assim sendo, esses dados são explorados para formulação das hipóteses e dos objetivos. Por fim, são apresentados indicadores com base nas hipóteses e objetivos.

Os cenários descritos incluem três tamanhos de base de dados com simulações de mudanças abruptas e graduais, bem como sete conjuntos de dados do mundo real. É importante destacar que, para cada conjunto de dados, na mesma linha, o melhor resultado é mostrado em **negrito**. Há também o uso de cores em algumas tabelas de resultados (mapa de calor) para destacar e exibir os dados comparativos em que os valores são representados com padrões de cores no fundo da célula. O uso do mapa de calor é para ajudar a transmitir informações instantaneamente, para isso, são combinadas duas cores principais e intensidade que variam em dois tons nas cores, verde e vermelho para indicar resultados bons e ruins, respectivamente. O verde mais escuro indica os maiores valores, em contraste o vermelho mais escuro indicado os menores valores. Esse formato de tabela complementado por cores fornece uma visão geral de dados fácil de entender.

3.1.2 Exploração dos resultados

A exploração dos resultados é a segunda fase e consiste na decomposição e padronização desses resultados. A organização é feita procurando categorizar e quantificar valores, possibilitando interpretações e inferências. Esta é a uma tarefa analítica.

3.1.3 Tratamento dos resultados, inferência e interpretação

Esta etapa é destinada ao tratamento dos resultados; ocorre nela o resumo e o destaque das informações para análise, levando as interpretações inferenciais.

Após a realização dos experimentos, as medidas de interesse são coletadas e organizadas em planilhas de texto. Para avaliar o desempenho dos detectores foi utilizado o teste *Wilcoxon Signed-rank* (DEMŠAR, 2006; BENAVALI; CORANI; MANGILI, 2016) com 95% de intervalo de confiança ($1-\alpha$) para mediana das diferenças, ou seja, $\alpha = 5\%$ de nível de significância e $1 - \beta$ como poder estatístico. Nos casos em que as análises rejeitam a hipótese nula entre pares de tratamentos, evidencia-se a superioridade de um tratamento em relação ao outro. Por fim, esse procedimento estatístico foi adotado devido à sensibilidade, indicações de uso na literatura da área e por ser um teste que não leva em consideração a quantidade de tratamentos.

É importante destacar que a avaliação de desempenho contribui para aferir o quão bons são os métodos detectores nos cenários adotados neste trabalho. O desempenho dos métodos é avaliado por meio de critérios de desempenho e técnicas da estatística que avaliam três hipóteses, no qual H_0 pressupõe igualdade estatística entre os métodos; H_1 pressupõe que \bar{X}_1 é superior estatisticamente à \bar{X}_2 ; e por fim, H_2 pressupõe que \bar{X}_1 é inferior estatisticamente à \bar{X}_2 .

Sendo assim, os métodos detectores de melhor desempenho devem oferecer um compromisso justo entre vitórias (superioridades estatísticas) e derrotas (inferioridades estatísticas). Nesses termos de vitórias e derrotas, os empates (igualdades estatísticas) são indiferentes para fins práticos da avaliação. Logo, a classificação ordenada dos métodos detectores é estabelecida pela diferença entre $\sum H_1 - \sum H_2$ para cada um dos três critérios avaliados e método de aprendizado.

Todos os três critérios são considerados na avaliação e notas são atribuídas para cada detector, considerando a diferença entre superioridades e inferioridades estatísticas $\sum H_1 - \sum H_2$. Assim que a avaliação empírica for concluída, a comparação dos métodos detectores pode ser iniciada.

O critério de seleção dos detectores será uma combinação entre métodos de aprendizado e medidas de desempenho acurácia da classificação, MTR e MCC que resultam em um índice final de seleção. Não há ponderação diferente entre as métricas para cálculo da média. Para cada critério, são assumidos pesos iguais de importância, ou seja, média aritmética das diferenças calculadas.

Para a obtenção de resultados fidedignos, foram empregadas duas abordagens metodológicas, por meio da perspectiva de HT e NB. Os resultados são apresentados numa sequência de quatro partes. A primeira compreende a análise das médias da métrica da acurácia da classificação. Na etapa seguinte, procede-se a avaliação das médias da métrica MTR. Na sequência, os dados são analisados sob a perspectiva das

médias da métrica MCC. Finalizando, conclui-se comparando os resultados obtidos a partir das duas perspectivas: médias das diferenças estatísticas calculadas e médias dos postos (de modo crescente - maior valor absoluto resulta em maior posto). Para isso, são identificados os aspectos gerais dos métodos, bem como os resultados semelhantes e discordantes apresentados pelos métodos de análise. Apresentam-se as médias dos postos para verificação dos resultados e constatação se corroboram ou não corroboram com a análise das diferenças estatísticas. Para obter essas médias, calculam-se os postos de cada método para cada métrica nos dois classificadores, seguidamente, obtém-se a média dos postos.

Finalmente, uma lista com a classificação dos métodos é definida e serve como referência para a decisão de quais detectores devem ser comparados com os métodos propostos na tese. Essa decisão é uma alternativa para reduzir a quantidade de dados apresentados. Vale ressaltar que essa seleção é baseada em técnicas estatísticas. A seguir, são apresentados mais detalhes sobre a avaliação empírica.

3.2 AVALIAÇÃO EMPÍRICA

Esta seção descreve as informações relevantes sobre os experimentos realizados como parte desta tese. Assim sendo, são apresentadas as particularidades sobre estações de trabalho, software, algoritmos de aprendizado, métodos de detecção de MC, geradores de dados artificiais e conjuntos de dados reais adotados nas experimentações.

3.2.1 Estação de trabalho

Os experimentos foram executados utilizando duas estações de trabalho. A Tabela 3 apresenta as principais informações sobre elas.

Tabela 3 – Resumo técnico das estações de trabalho utilizadas para realização dos experimentos.

| # | Sistema Operac. | Processador | Memória Princ. | Memória Secun. |
|---|--------------------|-----------------|----------------|----------------|
| 1 | Ubuntu 18.1/64-bit | i9-9900K/3.6GHz | 32GB/2666MHz | 250GB/SSD |
| 2 | Ubuntu 18.1/64-bit | i7-8700K/4.7GHz | 16GB/2333MHz | 250GB/SSD/NVMe |

Fonte: O AUTOR (2020).

3.2.2 Execução de experimentos

Todos os experimentos realizados para avaliar o desempenho dos métodos foram escritos/adotados e executados por meio do *Massive Online Analysis (MOA) framework/2014*. Além disso, também foram utilizados métodos detectores, geradores de dados artificiais, conjuntos de dados reais e metodologia para avaliação de desempenho

em ambiente de fluxo de dados *Prequential* (DAWID, 1984). No entanto, há três variações mais comuns: *Basic Window* (BW), *Sliding Window* (SW) e *Fading Factors* (FF) (GAMA; SEBASTIÃO; RODRIGUES, 2013). Foi adotada a variação SW por ter se apresentado como estratégia mais apropriada a ser utilizada em avaliações experimentais para estimar a acurácia de classificação em fluxo de dados com mudanças de conceito (HIDALGO; MACIEL; BARROS, 2019).

O MOA possui duas interfaces para acesso: uma baseada em interface gráfica e outra de linha de comando. Ambas interfaces são compatíveis com os scripts de parametrização de comandos do MOA, ou simplesmente chamados de scripts MOA. Para complementar os recursos disponíveis no MOA, algumas melhorias foram desenvolvidas (apresentadas na seção 6).

Para gerenciar as execuções dos experimentos no MOA, foi desenvolvida e adotada a ferramenta *Massive Online Analysis Manager PHP framework* (MOAManager) (MACIEL; SANTOS; BARROS, 2020). Além disso, ela também foi utilizada para extração e análise dos resultados.

3.2.3 Geradores e bases de dados reais

Esta subseção apresenta os detalhes sobre os geradores de dados artificiais e reais escolhidos para os experimentos utilizados nesta tese. Foram selecionados seis geradores de dados artificiais e parametrizadas duas versões de velocidade de mudança de conceito, representando mudanças abruptas e graduais. Cada gerador foi configurado com quatro MC, distribuídos em intervalos regulares e três tamanhos. Por fim, foram adotadas sete bases de dados reais, a saber: *Airlines*, *Connect4*, *Outdoor*, *Rialto*, *SpamData*, *WineRede* e *WineWhite*, apresentadas na seção 2.

Para simular as mudanças abruptas, diferentes conceitos foram simplesmente unidos. No caso de mudanças graduais, uma função de probabilidade disponível no MOA é utilizada para aumentar as chances de selecionar instâncias do novo conceito em vez de instâncias do anterior, e os dois conceitos coexistem durante a transição (graduação) de 500 instâncias.

Para cada gerador, mantiveram-se as parametrizações de funções de criação de conceitos, independentemente da velocidade da MC e tamanho dos conceitos. Vale ressaltar que para simular as foram concatenadas funções de geração de conceitos diferentes. Por exemplo: conceito 1 - vai da instância 1 até 4000; conceito 2 - vai da instância 4001 até 8000 em uma base de 8000 instâncias com uma (1) MC abrupta. Por fim, para cada parametrização (*script* MOA), foram executados 40 repetições, variando as sementes de inicialização de 1 a 40.

No gerador *Agrawal*, os conceitos são definidos por meio de funções de conceitos. Há dez funções de conceitos disponíveis no MOA: F1, F2, F3, F4, F5, F6, F7, F8, F9 e F10. Em um conjunto de dados com quatro MC, há cinco conceitos. Foram adotadas e concatenadas

para simular as funções de geração de conceitos F5, F6, F7, F8 e F9, respectivamente. Por fim, os demais parâmetros do gerador foram adotados os valores padrões, com exceção do parâmetro que define a velocidade da mudança. Isso vale também para os demais geradores, apresentados no próximo parágrafo.

Em LED, os conceitos também são definidos por meio de funções de conceitos. Há oito funções de conceitos disponíveis no MOA para ele: D0, D1, D2, D3, D4, D5, D6 e D7. Foram adotadas e concatenadas para simular as funções de geração de conceitos D1, D2, D3, D4 e D5, respectivamente. Na prática, essas funções simplesmente modificam a posição das barras de LED para simular MC.

Assim como na *Agrawal* e LED, o *Mixed* tem os conceitos definidos por meio de funções de conceitos. Há duas funções de conceitos disponíveis no MOA para ele: F1 e F2. Foram adotadas e concatenadas para simular as funções de geração de conceitos F1, F2, F1, F2 e F1, respectivamente.

Diferentemente dos geradores anteriores, o RRBF tem os conceitos definidos mudando o número de centroides. Há muitas possibilidades de parametrizações para criação de conceitos. Assim sendo, foram adotadas e concatenadas para simular as parametrizações fixas de duas classes, cinco atributos, vinte centroides e vinte centroides com *drift*, sendo os conceitos definidos variando as sementes de inicialização de cada conceito. Dessa forma, as sementes utilizadas foram 1, 2, 3, 4 e 5, respectivamente.

O gerador *Sine* também tem os conceitos definidos por meio de funções de conceitos. Há quatro funções de conceitos disponíveis no MOA: F1, F2, F3 e F4. Foram adotadas e concatenadas para simular as funções de geração de conceitos F1, F2, F3, F4 e F1, respectivamente.

Por último, o gerador *Waveform* é apresentado e, semelhante ao RRBF, não possui as funções de conceito bem definidas. Assim, os conceitos podem ser simulados variando a quantidade de atributos, sendo adotadas as quantidades 1, 5, 10, 15 e 20, respectivamente. Ao mudar a quantidade de atributos as posições que simulam os conceitos mudam. Adicionalmente, uma parametrização que gera ruído na base foi ativada.

3.2.4 Procedimentos de coleta de dados

Após execução bem sucedida dos experimentos usando a ferramenta MOAManager, cada *script* MOA gera um arquivo de resultados. As métricas utilizadas nas avaliações são extraídas desses arquivos. Para avaliar o desempenho do algoritmo de aprendizado é utilizada a medida de acurácia da classificação. Os métodos de DMC são avaliados utilizando medidas de FN, FP e cálculos com base nessas medidas. Além disso, são avaliados também por meio da medida MTR.

3.3 CONSIDERAÇÕES FINAIS

O objetivo desta seção foi descrever os aspectos metodológicos desta tese. Assim sendo, como o protocolo de pesquisa, a coleta de dados, a análise de dados e discussão dos resultados. O protocolo contém os procedimentos, parâmetros e as regras gerais que foram seguidas. Ele é uma das principais maneiras para aumentar a confiabilidade da pesquisa e destina-se a orientar pesquisadores ao conduzir uma reprodução dos experimentos.

Para que os resultados experimentais se tornem de fácil leitura e interpretação, procurou-se elaborá-los de uma maneira atraente, considerando quatro pontos de vista: (a) ser claro; (b) ser completo; (c) considerar alternativas; (d) mostrar evidências empíricas suficientes. Para alcançar os objetivos das avaliações experimentais, será seguida uma apresentação lógica dos resultados, com estilo direto, claro e minimizando ambiguidades.

Por fim, foram utilizados vinte detectores para que se proporcionasse uma aproximação mais adequada e/ou abrangente ao tema estudado. Na área de MC não se nega nem se minimiza a importância da coexistência da necessidade de vários métodos detectores para análise de desempenho, contudo, a coerência entre eles e a avaliação de desempenho é premissa básica.

4 DETECTORES DE MÚLTIPLA ANÁLISE

No apêndice C, foram avaliados os métodos de comitês de detectores do estado da arte em MC utilizando três formações de comitês: C_1 , C_2 e C_3 . Os resultados mostram que as melhores versões de formação de comitês para cada método foram AD com C_3 , ALHD com C_3 , ALO com C_3 e e -Detector com C_2 . Esses resultados são utilizados para comparar com as melhores versões dos métodos propostos e detectores adotados para compor os comitês, a saber: FHDDM, FSDD, FTDD e HDDMA. Resultados experimentais mostram que os métodos propostos alcançam resultados competitivos em comparação com o estado da arte.

Diversas abordagens foram propostas para aprendizado em fluxos contínuo de dados com mudança de conceito. Neste trabalho é adotada a abordagem de aprendizado adaptativo com detector embutido para realizar classificação em fluxo de dados. Os detectores geralmente são algoritmos simples e de crescimento linear com complexidade $\Theta(1)$. Para realização das detecções, esses algoritmos se concentram na análise de alterações na distribuição de dados, levando em consideração os acertos e os erros da classificação. Além disso, muitos deles usam um classificador auxiliar que antecipam treinamento antes da mudança de conceito para utilizá-lo após a mudança, substituindo o classificador principal pelo auxiliar.

Outras propostas agrupam classificadores usando diferentes funções de ponderação para calcular a classificação resultante, a exemplo do *Dynamic Weighted Majority* (DWM) (KOLTER; MALOOF, 2007), ARF (GOMES et al., 2019) e *Kappa Updated Ensemble* (KUE) (CANO; KRAWCZYK, 2019), enquanto algumas se concentram no problema de lidar com conceitos recorrentes (GONÇALVES JR.; BARROS, 2013a). Muitos desses trabalhos de comitês também adotam detectores como algoritmos auxiliares ao trabalho de classificação (BARROS, 2017; GONÇALVES JR.; BARROS, 2013b; BIFET, 2009; BIFET et al., 2010; MINKU; YAO, 2012; SANTOS et al., 2014).

Um número considerável de métodos de DMC foi proposto nos últimos anos. Exemplos incluem: DDM (GAMA et al., 2004), EDDM (BAENA-GARCIA et al., 2006), ADWIN (BIFET; GAVALDÀ, 2007a), ECDD (ROSS et al., 2012), STEPDP (NISHIDA; YAMAUCHI, 2007), HDDM (FRÍAS-BLANCO et al., 2015), etc.

Esta Seção 4 apresenta e descreve detectores de múltipla análise (comitês de detectores) para otimizar falsos alarmes de detecção. Os comitês são usados de forma semelhante aos detectores convencionais de única análise, trabalhando em conjunto com um método de aprendizado a partir de fluxos de dados e que pode ser baseado em um único classificador ou em múltiplos. Até certo ponto, a natureza da proposta dos métodos de múltipla análise desta tese é semelhante à do e -Detector (DU et al., 2014).

Os comitês de referência definidos no apêndice C também são utilizados nesta seção.

Entretanto, são utilizadas parametrizações diferentes na expectativa de fornecer uma visão mais ampla dos métodos de comitês propostos. Assim sendo, dado o conjunto de métodos $M = \{\text{DDE}, \text{Mixed Signals Drift Detection Ensemble (MSDDE)} \text{ e } \text{Match Drift Detection Ensemble (MDDE)}\}$, no DDE há três variações (a, b e c) no parâmetro que define o número mínimo de membros do comitês em MC para que o comitê assuma que ocorreu uma MC. Para o MSDDE, também há três variações (a, b e c) no parâmetro que define o número mínimo de membros do comitês em MC e *warning* para que o comitê assuma que ocorreu uma MC. Para exemplificar os experimentos com DDE e MSDDE, para ambos os métodos, os comitês $C_1 = (\text{FHDDM}, \text{FSDD} \text{ e } \text{FTDD})$ e $C_2 = (\text{FHDDM}, \text{FSDD} \text{ e } \text{HDDMA})$ foram parametrizados com detecções de MC com três sensibilidades, 1, 2 e 3. No caso do comitê $C_3 = (\text{FHDDM} \text{ e } \text{FSDD})$, como só tem dois métodos, foram parametrizadas duas sensibilidades, 1 e 2.

Seguidamente, no caso do MDDE, o algoritmo além do estado da detecção, também é avaliada a semelhança entre detecções dos detectores membros do comitê. Sendo assim, foram parametrizadas duas variações para C_1 e C_2 , com parâmetros a e b , utilizados para definir o número mínimo de membros do comitês em MC que possuem semelhança entre as detecções para que o comitê assuma que ocorreu uma MC e há apenas uma parametrização para a versão do comitê C_3 , exigindo apenas uma combinação entre os dois detectores membros.

Em resumo, as variações são apresentadas a seguir:

- DDE: $\{C_{1a}, C_{1b}, C_{1c}; C_{2a}, C_{2b}, C_{2c}; C_{3a}, C_{3b}\}$
- MSDDE: $\{C_{1a}, C_{1b}, C_{1c}; C_{2a}, C_{2b}, C_{2c}; C_{3a}, C_{3b}\}$
- MDDE: $\{C_{1a}, C_{1b}; C_{2a}, C_{2b}; C_{3a}\}$

As letras a, b e c nos nomes dos comitês foram utilizadas para diferenciar as versões de parametrizações. Essas letras substituem a progressão aritmética linear de 1 a 3, que indica o número mínimo de sensibilidade para que ocorra uma DMC. Vale ressaltar que o método proposto MDDE com o comitê C_3 foi testado com uma combinação e os comitês C_1 e C_2 com duas combinações.

No apêndice D, as Tabelas 62, 63, 68, 69, 74 e 75 apresentam os resultados dos testes das hipóteses estatísticas (aplicando o teste de *Wilcoxon Signed-Rank*) utilizando as médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. As Tabelas 64, 65, 70, 71, 76 e 77 apresentam os resultados dos testes das hipóteses estatísticas com as médias de MTR dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Seguidamente, nas Tabelas 66, 67, 72, 73, 78 e 79 são apresentados os resultados dos testes das hipóteses estatísticas (aplicando o teste de *Wilcoxon Signed-Rank*) com as médias de MCC dos métodos testados em todos os

conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. É importante observar que, para cada uma das três hipóteses H_0 , H_1 e H_2 testadas, na mesma linha, o melhor resultado é mostrado em **negrito**. Essas tabelas apresentam a informação de qual método foi igual (empate), superior (vitória) ou inferior (derrota) estatisticamente a qualquer método avaliado. Na avaliação as derrotas são subtraídas das vitórias para obter o rank (das diferenças).

Ainda no apêndice D, a Seção D.1 apresenta os resultados e análise das avaliações realizadas com diferentes parametrizações e os algoritmos DDE, MSDDE e MDDE. É importante destacar que, para cada um dos três algoritmos, foram realizadas avaliações de desempenho individuais por algoritmo, objetivando identificar qual parametrização resultou em melhores resultados. Por fim, como resultado dessa avaliação foram selecionados os parâmetros adotados nos métodos propostos e utilizados na Seção 4.6.

4.1 MÚLTIPLA DETECÇÃO

A essência da detecção de múltipla análise é tomar decisões baseadas em outras avaliações. Esses métodos são chamados de ensemble (do Inglês, traduzindo para o português seria comitê) e são projetados para avaliar alternativas e tomar decisões sobre uma ampla gama de problemas. As alternativas são fornecidas pelos membros do comitê e estes são chamados em mineração de dados de especialistas. Sendo assim, há diferentes estratégias de tomada de decisão de combinação dessas alternativas e cada uma envolve incerteza e risco.

A estratégia de decisão também pode envolver análises qualitativas e quantitativas, e algumas delas são bem sucedidas usando uma forma de análise sobre a outra. Tomadas de decisões podem ser afetadas não apenas pela estratégia de combinação, mas também pela análise dos membros do comitê. Na prática, decidir fornece uma posição em relação ao futuro.

No caso da DMC, as decisões são tomadas sem garantias da correteza, pois não há uma supervisão que indique os erros de decisão. A estratégia de comitê de detecção precisa agir sem saber com certeza todas as consequências de suas decisões. A incerteza influencia no aumento do número de resultados possíveis, e as consequências desses resultados precisam ser consideradas.

É importante para a estratégia de tomada de decisão identificar quais são as incertezas, quais são os resultados possíveis e quais serão as consequências. A seleção dos membros dos comitês às vezes pode esclarecer o problema em que estão trabalhando, sendo possível escolher membros com maiores probabilidades a cada resultado possível.

A aversão à probabilidade de insucesso (ou, simplesmente, risco) é um fator que afeta uma boa tomada de decisão. É preciso levar em consideração que detectores que exibem comportamento individual de baixo desempenho se tornarão candidatos a risco quando

utilizados em um comitê. Por exemplo, em uma partida de futebol a maioria das pessoas irá selecionar para a sua equipe os jogadores com melhor reputação, minimizando os riscos de uma certa perda.

Experimentalmente, tem sido mostrado que comitês de classificadores apresentam melhores desempenhos do que um único classificador (sistema decisório único - única análise). Os trabalhos (DU et al., 2014; MACIEL; SANTOS; BARROS, 2015) mostraram as primeiras estratégias de comitês para DMC disponíveis na literatura. Apesar de obterem melhores resultados, esses métodos apresentam uma maior complexidade de tempo para realização das tarefas.

A eficiência dos algoritmos é importante na programação competitiva. Assim sendo, geralmente os detectores são de complexidade $\Theta(1)$, como DDM, EDDM e STEPD. O tempo de execução de um algoritmo $\Theta(1)$ é constante, logo, não depende do tamanho da entrada. É comum em projetos de algoritmos encontrar uma solução que resolva um problema consumindo mais recursos (é possível mencionar o caso das soluções por comitês de classificadores), sendo desafiador inventar um algoritmo rápido. Os benefícios da solução em termos de resolução de problema podem justificar o uso de um algoritmo de maior complexidade.

Um algoritmo de complexidade linear passa pela entrada um número constante de vezes. Geralmente, essa é a melhor complexidade de tempo possível, porque geralmente é necessário acessar cada elemento de entrada pelo menos uma vez antes de ter um resultado. A complexidade de um algoritmo estima de quanto tempo ele necessitará para uma dada entrada. Assim sendo, a ideia é representar a eficiência como uma função na qual o parâmetro é dado pelo tamanho da entrada. Logo, ao calcular a complexidade do tempo, é possível saber se o algoritmo é rápido o suficiente sem conhecer detalhes de sua implementação.

Um comitê de detectores simples é de complexidade de tempo $\Theta(n)$ por fazer múltiplas análises de detecções. A função é Θ e n é a variável que indica o tamanho da entrada. Por exemplo, se a entrada for uma lista de detectores, n será o número de elementos da lista. O cálculo da complexidade de tempo não diz o número exato de vezes que um bloco de código é executado, mas mostra a ordem de magnitude. Logo, em termos de complexidade de tempo, um comitê de detectores é mais lento quando comparado a um algoritmo de $\Theta(1)$, pois a ideia básica de um comitê é utilizar mais de um membro, ou seja, n será maior que 1. Vale ressaltar que para os comitês utilizados na tese, foram adotados n no máximo igual a três.

Embora a complexidade de tempo de um comitê de detectores (múltipla análise) seja maior que um detector (única análise), há outros componentes envolvidos na estratégia de AADE que podem degradar o tempo e/ou acurácia. Por exemplo, um detector que sinalize muitos falsos alarmes pode comprometer a eficiência da estratégia de aprendizado, pois a estratégia vai precisar reconstruir o modelo a cada novo falso alarme, incrementando ao

tempo total o tempo dessa reconstrução. Além disso, o espaço amostral das previsões será minimizado, e como uma possível consequência uma menor acurácia da classificação será obtida. Assim, apesar de um comitê de detectores representar uma maior complexidade de tempo quando comparado à detecção de análise linear, do ponto de vista de análise de detecção, a redução de falsos alarmes de um detector é um atributo de qualidade justificável para escolha de um detector em um dado problema. A seguir, são apresentadas três estratégias para detecção de múltipla análise: uma delas é baseada no DDE (MACIEL; SANTOS; BARROS, 2015) e as demais não foram publicadas.

4.2 MOTIVAÇÃO

Com a crescente tendência em armazenamento, processamento e sumarização de dados, diferentes domínios de aplicações tais como reconhecimento de padrões, aprendizado de máquina, mineração de dados e mudança de conceito têm desenvolvido técnicas e métodos buscando automatizar a extração de informação dos dados. Nesses domínios os dados são essenciais para o progresso tecnológico. No caso da mudança de conceito, uma metodologia que é bastante conhecida neste tipo de pesquisa é de DMC, à qual uma grande quantidade de dados é exposta para o algoritmo de aprendizado e seu desempenho é monitorado por um detector que analisa em tempo real o comportamento desse método. Experimentos com detectores são capazes de monitorar grandes quantidade de dados e simular múltiplos cenários.

Técnicas de DMC são feitas para auxiliar no monitoramento do classificador, servindo como uma ferramenta para diagnóstico da MC (GAMA et al., 2004). Um detector pode ser classificado em várias categorias (GAMA et al., 2014), dependendo da forma de análise que ele contém. Por exemplo, os detectores DDM e STEPDP são métodos de análise sequencial e estatística, respectivamente. Recentemente, há iniciativas como Du et al. (2014), Maciel, Santos e Barros (2015), Wozniak et al. (2016) que combinam resultados de múltiplos detectores para tomada de decisão. Uma tarefa desafiadora para o comitê de detectores é como combinar e comparar informações a partir de métodos detectores para fornecer resultados mais precisos.

Atualmente, comitês com estratégia de votação é o mais frequentemente utilizado para obter informações combinando detectores (DU et al., 2014; MACIEL; SANTOS; BARROS, 2015; WOZNIK et al., 2016). Eles geralmente baseiam-se na similaridade entre os dados de detecção. Mais especificamente, o método de combinação por votação majoritária é o mais utilizado. Ele é conhecido por ser simples de programar e rápido na execução. Entretanto, é feita uma suposição de que todos os detectores de análise linear detectam a MC quase que simultaneamente, tornando-o inadequado em muitos problemas reais onde os dados possuem diferentes dispersões e frequências. Por fim, um ponto em aberto em MC é como combinar os alertas de detecções e identificar como eles contribuem para

um grau de semelhança entre diferentes análises de detecções. Dessa forma, a proposta é utilizar a combinação de alarmes de detecções para reduzir falsos alarmes.

4.3 DDE: DRIFT DETECTION ENSEMBLE

Esta subseção fornece uma descrição detalhada do DDE (MACIEL; SANTOS; BARROS, 2015), um comitê leve de DMC. São incluídas também a motivação e hipótese, bem como os detalhes do algoritmo. A principal ideia por trás do DDE é controlar a execução de diferentes detectores para melhorar a precisão e faz isso combinando os resultados individuais de detecção, adotando uma estratégia algébrica e nível de sensibilidade. O DDE compartilha com os detectores membros do comitê os resultados de acertos e erros das previsões do classificador. Por fim, o termo leve empregado anteriormente diz respeito à complexidade de tempo do DDE ser $\Theta(n)$ e de usar um único classificador.

O DDE foi experimentalmente avaliado em conjunto com a estratégia de AADE. Vale ressaltar que, para fins práticos, adotou-se o algoritmo de aprendizado composto por apenas um (1) classificador. Os resultados de acertos e erros do classificador são repassados ao DDE e encaminhados aos membros do comitê, desde que o membro não esteja classificado com *drift* (MC) pelo DDE ou a condição para o estado de *drift* não tenha ocorrido.

A estratégia para tomada de decisão refere-se à sensibilidade do comitê em relação às previsões dos seus membros, isto é, os detectores individuais. A ideia é contabilizar quais membros sinalizaram *warning* e *drift* a cada iteração. Tratando-se de detecção, será pouco provável que dois detectores sinalizem *drift* no mesmo momento (exatamente na mesma instância) e pensando nisso foi necessário adotar uma estratégia para normalizar essa situação. Assim, durante um determinado período de análises, o estado de *drift* será mantido, ficando o detector membro sem receber novos dados para análise. Por fim, no algoritmo publicado do DDE em (MACIEL; SANTOS; BARROS, 2015), foi utilizado um período de 100 análises (que corresponde a variável *maxWait* no algoritmo 1) e, posteriormente após novos experimentos, o valor 300 foi adotado por contribuir com a melhoria dos resultados.

Os critérios para tomada de decisão pelo comitê são baseados na soma dos sinais de *warnings* e *drifts* de cada um dos membros. A condição que indica o estado de *warning* pelo comitê é atribuída quando: soma de estados de *warnings* + soma de estados de *drifts* for igual ou maior ao nível de sensibilidade. Vale ressaltar que essas somas são baseadas nos detectores membros do comitê a cada análise de dados pelos detectores membros. O *drift* ocorre quando a soma de estados de *drifts* for igual ou maior ao nível de sensibilidade. Em caso de nenhuma das duas condições anteriores for verdade (ou seja, condição para *warning* ou *drift*), o fluxo é sinalizado como estável.

A principal motivação para uso do DDE é reduzir os falsos alarmes e contribuir com

a melhoria da acurácia *Prequential*. A hipótese para essa redução está na combinação de mais de um sinal de *drift*. Neste sentido, é indicado o uso de comitês com sensibilidade maior que 1, pois são necessários ao menos dois detectores para confirmar o sinal de *drift*. As parametrizações de sensibilidade podem ser alteradas de acordo com: objetivo esperado, detectores membros e/ou domínio do problema.

Uma vez que o comitê toma a decisão de *drift*, os membros são reinicializados. Essa medida é necessária para garantir que dados antigos não sejam levados em consideração na análise do futuro fluxo de dados. Por fim, as estatísticas de *warning* e *drift* do comitê também são reinicializadas.

O algoritmo 1 apresenta um pseudo-código do comitê que foi implementado em linguagem de programação Java. Em comparação ao algoritmo apresentado em (MACIEL; SANTOS; BARROS, 2015), apesar de existir algumas diferenças entre as versões, a técnica de detecção não mudou. Na primeira versão foram embutidos ao algoritmo os detectores membros. Houve também alteração no parâmetro da quantidade de instância para aguardar detecções, que foi alterado de $w = 100$ para $w = 300$. Essa decisão foi tomada de acordo com resultados experimentais, que mostraram melhoria na acurácia *Prequential*. Entre as linhas 32 e 44 são implementados os critérios para tomadas de decisões. O primeiro, na linha 32, avalia a condição de estabilidade do fluxo. Caso não seja atendida, a condição da linha 36 será avaliada, verificando o sinal de *warning*. Caso também não seja atendida, um *drift* é acionado.

Entre as linhas 6 e 31, são repassados os dados aos detectores membros do comitê e contabilizados seus sinais de *warnings* e *drifts*. A linha 6 possui uma estrutura de repetição na qual a coleção de dados é compartilhada com os detectores membros do comitê. Cada detector recebe a mesma resposta sobre acerto ou erro da predição da classificação, caso não esteja sinalizado pelo comitê com sinal de *drift*. Na linha 27, há uma condição de parada de análise de dados pelos detectores. Sempre que a condição for atendida, os demais detectores da coleção não recebem a predição do classificador, pois será desnecessário analisar, uma vez que a condição da linha 27 atende à condição para que haja *drift*.

Por fim, há outra diferença entre os algoritmos do DDE (MACIEL; SANTOS; BARROS, 2015) e o algoritmo 1. Inicialmente, o DDE não tinha capacidade de generalizar os membros do comitê, apenas era possível escolher um entre três comitês disponíveis. Além disso, também foi reescrita a parte de análise dos resultados do membros do comitê, período de espera e tomada de decisão. Entretanto, vale ressaltar que a estratégia de detecção manteve-se, ou seja, soma dos alertas e níveis de sensibilidades.

Algoritmo 1: DDE - DRIFT DETECTION ENSEMBLE**Input:**

stream: data stream of examples, *sens*: sensibility detection, *MaxWait*: max wait detection, *members*: members of detection

```

1 resp [*] ← 0
2 foreach instance in stream do
3   pred ← prediction (instance) //Prediction result of the classifier is received
4   instNumber ← instNumber + 1
5   driftCount, warningCount, index ← 0
6   foreach changeDetector in members do
7     if resp [index] < 1 then
8       changeDetector.input (pred)
9       if changeDetector is DRIFT then
10        resp [index] ← instNumber
11        driftCount ← driftCount + 1
12      end
13    else
14      if changeDetector is WARNING then
15        warningCount ← warningCount + 1
16      end
17    end
18  end
19  else
20    if resp [index] + maxWait < instNumber then
21      resp [index] ← 0
22    end
23    else
24      driftCount ← driftCount + 1
25    end
26  end
27  if driftCount >= sens then
28    break
29  end
30  index ← index + 1
31 end
32 if warningCount + driftCount < sens then
33   raise stable
34 end
35 else
36   if driftCount < sens then
37     raise warning
38   end
39   else
40     reset members
41     resp [*] ← 0
42     raise drift
43   end
44 end
45 end

```

Fonte: O AUTOR (2020).

4.4 MSDDE - MIXED SIGNALS DRIFT DETECTION ENSEMBLE

Esta subseção fornece uma descrição detalhada do MSDDE, outra proposta de comitê leve de DMC. Como já mencionado no tópico 4.3, que apresenta o DDE, a principal estratégia dele é combinar resultados individuais adotando uma abordagem algébrica e por sensibilidade. O MSDDE também compartilha com os detectores membros do comitê os resultados de acertos e erros das predições do classificador. Por fim, o termo leve empregado anteriormente para o DDE, também se aplica ao MSDDE, pois a complexidade de tempo dele é $\Theta(n)$.

O MSDDE também foi experimentalmente avaliado em conjunto com a estratégia de AADE. Vale ressaltar que, para fins práticos, também adotou-se o algoritmo de aprendizado composto por apenas um (1) classificador. Os resultados de acertos e erros do classificador são repassados pelo comitê aos membros do comitê, desde que: 1) o membro não esteja classificado com *drift* pelo comitê; ou 2) quando a condição para o estado de *drift* não seja verdade.

A estratégia para tomada de decisão é baseada no número mínimo necessário de membros do comitê com previsões de *warning* e/ou *drift*. A ideia é contabilizar quais

membros sinalizaram os estados de *warning* e *drift* a cada iteração. Como mencionado anteriormente na subseção do DDE, quando se trata de detecção, será pouco provável que dois detectores sinalizem *drift* no mesmo momento. Dessa forma, a mesma estratégia do DDE para normalizar essa situação é adotada.

Há três diferenças básicas entre MSDDE e DDE. A primeira diz respeito à atribuição de peso diferentes aos estados de *warning* e *drift*, no qual o peso do *warning* é equivalente a 1/2 do peso do *drift*. A segunda é que não existem dois contadores separados, um para *drift* e outro para *warning*. Existe apenas um contador que mistura os dois sinais levando em consideração os pesos de cada estado. A terceira diferença básica está na condição dos estados de decisão, pois é levado em consideração o contador da mistura dos sinais para avaliar os níveis de *warning* e/ou *drift*. Com o objetivo de não deixar o texto repetitivo, o foco daqui para frente será nas diferenças entre os comitês DDE e MSDDE.

Os critérios para tomada de decisão pelo comitê são baseados na soma dos sinais de *warnings* e *drifts* dos membros do comitê. A condição que indica o estado de *warning* pelo comitê é atribuída quando: soma dos pesos de *warning* + soma dos pesos de *drift* maior que nível definido para *warning*. Vale ressaltar que essas somas são baseadas nos detectores membros do comitê e os estados de *warning* possuem peso igual a 1/2 do peso do *drift*. O *drift* ocorre quando a soma dos pesos desses estados for igual ou maior (nível definido para o *drift*). Caso nenhuma das duas situações ocorra *warning* ou *drift*, o fluxo é sinalizado como estável.

A principal motivação para uso desse comitê é reduzir o número de falsos alarmes. A hipótese para essa redução está na mistura dos sinais de *warning* e *drift* dos membros do comitê. Nesse sentido, também é indicada a escolha por comitês com nível mínimo maior que 1. As parametrizações de níveis de *warning* e *drift* podem ser alteradas de acordo com: 1) objetivo esperado; 2) detectores membros; e/ou 3) domínio do problema.

O algoritmo 2 apresenta um pseudo-código do comitê que foi implementado em linguagem de programação Java. O algoritmo do MSDDE é semelhante ao algoritmo do DDE. Entre as linhas 32 e 44 estão os critérios para tomadas de decisões em relação à classificação do fluxo de dados.

Semelhante ao algoritmo 1, na linha 32 é avaliada a condição de estabilidade do fluxo. Quando é falsa a condição de estabilidade, é avaliada a condição da linha 36, que verifica se a soma dos pesos é menor que o nível definido para *drift*, caso seja, é acionado o estado de *warning* e, caso também não seja atendida é acionado um *drift*.

Entre as linhas 6 e 31 são repassados os dados aos detectores membros do comitê e contabilizados seus sinais de *warnings* e *drifts*, levando em consideração seus respectivos pesos. A diferença entre os algoritmos MSDDE e DDE está apenas na atribuição dos pesos quando os membros do comitê sinalizam *warning* ou *drift* e modo de contabilizar os sinais. Vale ressaltar que o esquema de soma que mistura os sinais de *warning* e *drift* com seus respectivos pesos permite parametrizações sensíveis de ponto flutuante para os

níveis de sinais utilizados para tomadas de decisão. Devido a essa mistura dos sinais na soma, é possível, por exemplo, parametrizar um nível de *drift* como 1,5, ou seja, o peso de um *drift* e meio.

Algoritmo 2: MSDDE - MIXED SIGNALS DRIFT DETECTION ENSEMBLE

Input:

```

stream: data stream of examples, maxWait: max wait detection, members: members of ensemble
1 resp [*] ← 0
2 foreach instance in stream do
3   predRes ← prediction (instance)
4   instNumber ← instNumber + 1
5   level, index ← 0
6   foreach changeDetector in members do
7     if resp [index] < 1 then
8       changeDetector.input (predRes)
9       if changeDetector is DRIFT then
10        resp [index] ← instNumber
11        level ← level + 1
12      end
13    else
14      if changeDetector is WARNING then
15        level ← level + 1/2
16      end
17    end
18  end
19  else
20    if resp [index] + maxWait < instNumber then
21      resp [index] ← 0
22    end
23    else
24      level ← level + 1
25    end
26  end
27  if level >= driftLevel then
28    break
29  end
30  index ← index + 1
31 end
32 if level < warningLevel then
33   raise stable
34 end
35 else
36   if level < driftLevel then
37     raise warning
38   end
39   else
40     reset members
41     resp [*] ← 0
42     raise drift
43   end
44 end
45 end

```

Fonte: O AUTOR (2020).

4.5 MDDE - MATCH DRIFT DETECTION ENSEMBLE

Esta subseção fornece uma descrição detalhada do MDDE, outra proposta de comitê de DMC. Como já mencionado no DDE e MSDDE, a principal ideia dele também é combinar resultados individuais adotando uma estratégia de sensibilidade. O MDDE também compartilha com os detectores membros do comitê os resultados de acertos e erros das previsões do classificador. Por fim, o termo *leve* empregado anteriormente no DDE e MDDE, não se aplica a esta proposta, pois a complexidade do MDDE é $\Theta(n^2)$.

O MDDE também foi experimentalmente avaliado em conjunto com a estratégia de AADE. Vale ressaltar que, para fins práticos, também se adotou o algoritmo de aprendizado composto por apenas um (1) classificador. Os resultados de acertos e erros do classificador são repassados pelo comitê aos membros do comitê, desde que o membro não esteja classificado com *drift* pelo comitê ou a condição para o estado de *drift* não tenha sido atendida.

A estratégia para tomada de decisão é baseada em número mínimo de combinações necessárias de membros do comitê com previsões de *drift*. A ideia é combinar membros que sinalizaram *drift* usando o erro médio da classificação em um período de dados. Há duas estratégias parametrizáveis para tomada de decisão do *drift*: por similaridade ou dissimilaridade entre detecções. Como mencionado anteriormente na seção do DDE, tratando-se de detecção, será pouco provável que dois detectores sinalizem *drift* no mesmo momento e, portanto, é adotada a mesma estratégia do DDE para normalizar essa situação.

Os critérios para tomadas de decisões pelo comitê são baseados na soma dos sinais de *warnings* e *drifts* dos membros do comitê. A condição que indica o estado de *warning* pelo comitê é atribuída quando: soma de estados de *warnings* ou soma de estados de *drifts* for maior que zero. Vale ressaltar que essas somas são baseadas nos detectores membros do comitê a cada análise de dados pelos detectores membros. Uma DMC ocorre quando a soma de estados de MC for maior que 1 e ao menos dois detectores tenham combinação na regra escolhida, podendo ser similaridade ou dissimilaridade entre detecções para que ocorra a DMC. Caso nenhuma das situações anteriores seja verdadeira, o fluxo é sinalizado como estável.

Há três diferenças básicas entre MDDE e DDE. A primeira diz respeito ao modo como classificar *warning*. No MDDE basta qualquer membro do comitê ter sinalizado *warning* ou *drift* para ser classificado como estado de *warning*. A segunda diz respeito ao modo como classificar *drift*. Dessa forma, o *drift* não dependerá apenas de uma quantidade de detecções, mas também que estes atendam a condições da estratégia selecionada de similaridade ou dissimilaridade. Por fim, a classificação de estabilidade do fluxo, como consequência, também é diferente, bastando a condição para *warning* ser falsa. Para que o texto não fique repetitivo ao longo desta tese, será focado daqui para frente as diferenças entre os comitês DDE e MDDE.

Quando um detector membro diz que ocorreu MC, é registrado o percentual de erro médio da acurácia da classificação. Uma janela deslizante é utilizada para avaliar em tempo real o erro médio da classificação e esse erro médio calculado é utilizado como parâmetro de referência para avaliar a similaridade ou dissimilaridade entre detecções. Em caso no qual mais de um membro do comitê esteja com alerta de MC, será avaliada a regra parametrizada de similaridade ou dissimilaridade, e, então, tomada a decisão.

O algoritmo 3 apresenta um pseudo-código do comitê que também foi implementado em linguagem de programação Java. O algoritmo do MDDE é semelhante aos algoritmos do DDE e MSDDE. Nas linhas 35-54 são avaliadas as condições de semelhança de detecção. Na linha 38, é calculada a variável ***diff*** que diz respeito ao valor absoluto da diferença entre os erros médios por pares de detecções em MC. Em caso no qual a estratégia parametrizada seja combinação por dissimilaridade, será verificada a condição se ***diff*** é maior ou igual ao limite parametrizado para MC, chamado de ***driftLimit***. Na combinação

por similaridade, será verificada a condição se *diff* é menor ou igual *driftLimit*. Nos casos em que a condição é verdade, a variável *driftAlert* recebem mais 1 ao valor atual.

Uma DMC ocorre quando a quantidade de combinações for igual ao tamanho parametrizado de combinações, dado por *minMatch*. Apesar de ser possível todos os membros estarem sinalizados com alerta de MC e ser possível que todos atendam à regra de combinação, não é necessário fazer todas as combinações se o número mínimo de combinações for atendido antes disso.

A principal motivação para uso desse comitê é reduzir o número de falsos alarmes. A hipótese para essa redução está na combinação de dois detectores, seja adotando a estratégia de similaridade ou dissimilaridade. Nesse sentido, é indicada a escolha por comitês com uma combinação, independente da estratégia adotada. As parametrizações para DMC podem ser alteradas de acordo com o objetivo esperado.

Nas linhas 31-56 estão os critérios para tomadas de decisões em relação a classificação do fluxo de dados. Já na linha 31 é testada a condição de estabilidade do fluxo. Caso não seja atendida, a condição da linha 33 que verifica a ocorrência de detectores em estado de *drift* será avaliada. Caso seja verdade, serão avaliadas as similaridade de detecções. Em caso de confirmação da similaridade ou dissimilaridade, dependendo da parametrização, o estado do *drift* será acionado.

Entre as linhas 11-30 são repassados os acertos ou erros da classificação aos detectores membros do comitê e contabilizados seus alertas. Além disso, também é calculado o erro médio da acurácia no momento da DMC. A diferença entre os algoritmos MDDE e MSDDE está no modo utilizado para testar as condições do fluxo de dados, pois no MDDE não há distinção entre pesos para sinais de *warning* e *drift*, assim como também o modo para testar o condição de *drift* não leva em consideração apenas os sinais dos detectores membros.

Algoritmo 3: MDDE - MATCH DRIFT DETECTION ENSEMBLE**Input:**

stream: data stream of examples, *minMatch*: minimum matching detectors, *MaxWait*: max wait detection, *members*: members of detection, *w*: window size

```

1  resp [*], instNumber, numErr, pos, stPred [*] ← 0
2  foreach instance in stream do
3    pred ← prediction(instance)
4    instNumber ← instNumber + 1
5    numErr ← numErr - stPred[pos] + pred
6    stPred[pos] ← pred
7    pos ← pos + 1
8    if pos = w then
9      | pos ← 0
10   driftLimit, warningCount, index ← 0
11   foreach changeDetector in members do
12     if resp[index] < 1 then
13       changeDetector.input(pred)
14       if changeDetector is DRIFT_LEVEL then
15         resp[index] ← instNumber
16         driftLimit ← driftLimit + 1
17         err ← 0
18         if numErr > 0 then
19           | err ← numErr / w
20         stMeanErr[index] ← err
21       else
22         if changeDetector is WARNING_LEVEL then
23           | warningCount ← warningCount + 1
24       else
25         if resp[index] + maxWait < instNumber then
26           resp[index] ← 0
27         else
28           | driftLimit ← driftLimit + 1
29         index ← index + 1
30   end
31   if warningCount > 0 OR driftLimit > 0 then
32     raise warning
33     if driftLimit > 1 then
34       driftAlert, countMatch, detectors[*], n, m ← 0
35       for n ∈ {1, ..., membersTotal} do
36         for m ∈ {1, ..., membersTotal} do
37           if n <> m AND resp[n] <> 0 AND resp[m] <> 0 then
38             diff ← |(stMeanErr[n] - stMeanErr[m])|
39             if strategy = dissimilarity then
40               if diff >= driftLimit then
41                 | driftAlert ← 1
42             else
43               if diff <= driftLimit then
44                 | driftAlert ← 1
45             if driftAlert = 1 then
46               if not n ∈ detectors OR not m ∈ detectors then
47                 detectors.add(n, m)
48                 countMatch ← countMatch + 1
49                 if countMatch = minMatch then
50                   reset members
51                   resp[*], stPred[*], numErr, pos ← 0
52                   raise drift
53             end
54           end
55         else
56           raise stable
57   end

```

Fonte: O AUTOR (2020).

4.6 RESULTADOS E DISCUSSÃO

Esta subseção reúne os métodos DDE, MSDDE e MDDE, que obtiveram os melhores resultados nas variações de comitês, apresentado no apêndice D nas subseções D.1.1, D.1.2 e D.1.3. Além desses, também reúne os métodos FHDDM, FSDD, FTDD e HDDMA, utilizados para formar os comitês C_1 , C_2 e C_3 . Finalmente, também foram incluídos os métodos AD, ALHD, ALO e e -Detector que obtiveram os melhores resultados nas variações de comitês, apresentados no apêndice C. Por fim, o objetivo desta seção é avaliar os métodos de comitês propostos com outras abordagens semelhantes, incluindo também os métodos convencionais de detecção utilizados para formação dos comitês.

Para todos os métodos de comitês propostos nesta seção, a variação de comitê C_2 com os detectores FHDDM, FSDD e HDDMA proporcionou os melhores resultados. É importante destacar também que os melhores resultados foram alcançados com combinação de, no mínimo, dois detectores e foram avaliados três casos, a saber: (i) apenas uma detecção; (ii) duas detecções; e (iii) três detecções. Vale ressaltar que há diferença nas regras para combinar os sinais de detecções nos comitês propostos.

Para os métodos de comitês apresentados como estado da arte em detecção, a variação de comitê C_3 proporcionou os melhores resultados para os métodos AD, ALHD, ALO, entretanto, isso não se manteve no e -Detector, no qual a melhor variação foi com C_2 . Tratando-se dos métodos convencionais de DMC, os métodos utilizados para formar os comitês foram selecionados com base em critérios de desempenhos.

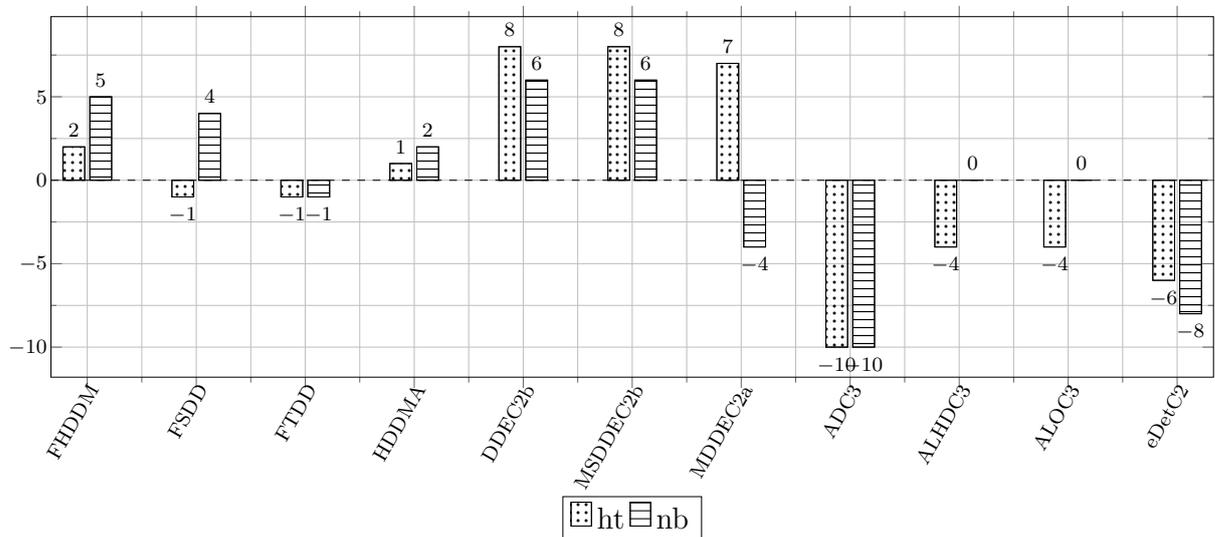
As Tabelas 4 e 5 apresentam os resultados das médias de acurácias dos métodos FHDDM, FSDD, FTDD, HDDMA, AD, ALHD, ALO, e -Detector, DDE, MSDDE e MDDE testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos que obtiveram os melhores resultados com o HT e NB foram MDDEC_{2a} (3, 87) e FSDD (4, 50), respectivamente. O recurso do mapa de calor mostra uma maior concentração de vermelho nas linhas dos métodos AD, ALHD, ALO e e -Detector com o HT e NB, implicando nos piores resultados.

A Figura 5 (a) apresenta o resumo dos resultados com HT e NB apresentados no apêndice D nas Tabelas 80 e 81. Em termos das hipóteses testadas, os métodos DDE.C_{2b} e MSDDE.C_{2b} com HT forneceram o maior número de vitórias (8) e em termos estatísticos foram iguais ao MDDEC_{2a} e superiores aos demais métodos. No caso do HT, é importante destacar que os comitês DDE.C_{2b} e MSDDE.C_{2b} superaram os detectores membros em termos de comparações estatísticas e apenas o FHDDM não foi inferior estatisticamente ao MDDEC_{2a}. Vale ressaltar que os membros FHDDM, FSDD, FTDD e HDDMA foram iguais estatisticamente. Já com NB os melhores em termos de quantidade de vitórias foram o DDE.C_{1b} e o MSDDE.C_{2b}, ambos com 6 vitórias. Em termos estatísticos eles foram iguais aos métodos FHDDM, FSDD e HDDMA. O método com maior número de derrotas usando o HT e NB foi ADC₃, com 10 derrotas, sendo inferior estatisticamente a todos os demais métodos.

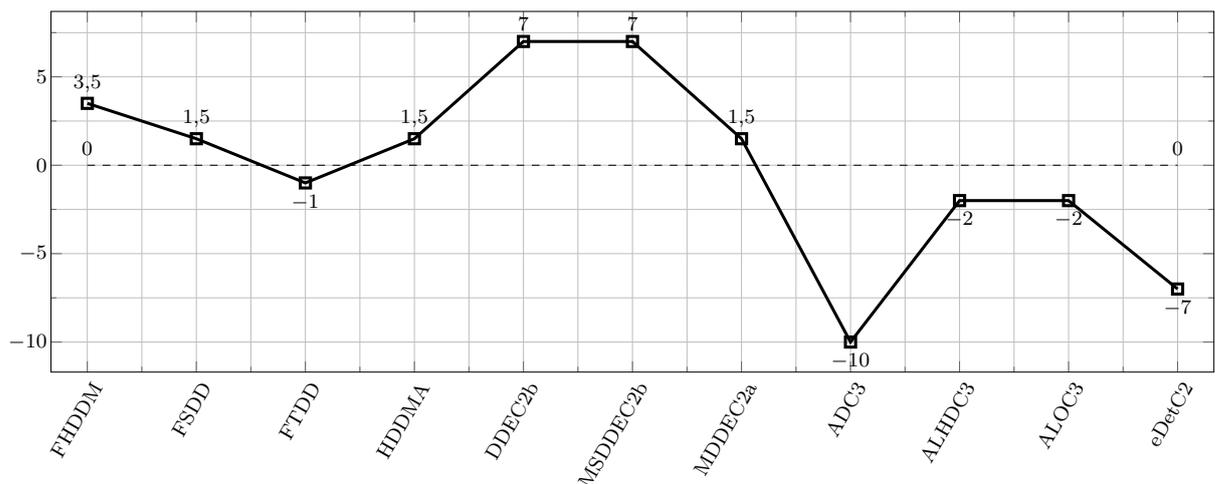
É possível notar que os melhores resultados nos conjuntos de dados avaliados foram obtidos pelos métodos DDE.C_{2b}, MSDDE.C_{2b}, FHDDM, HDDMA, FSDD e MDDEC_{2a} todos com diferenças calculadas acima de zero, como pode ser visto na Figura 5 (b). Apesar disso, vale ressaltar que os desempenhos dos métodos DDE.C_{2b} e MSDDE.C_{2b} foram relevantes na maioria das situações, e houve superioridade estatística em relação aos demais métodos com o HT, com exceção ao MDDEC_{2a}. A Figura 5 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que os métodos DDE.C_{2b} e MSDDE.C_{2b} obtiveram a maior média (7)

Figura 5 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.

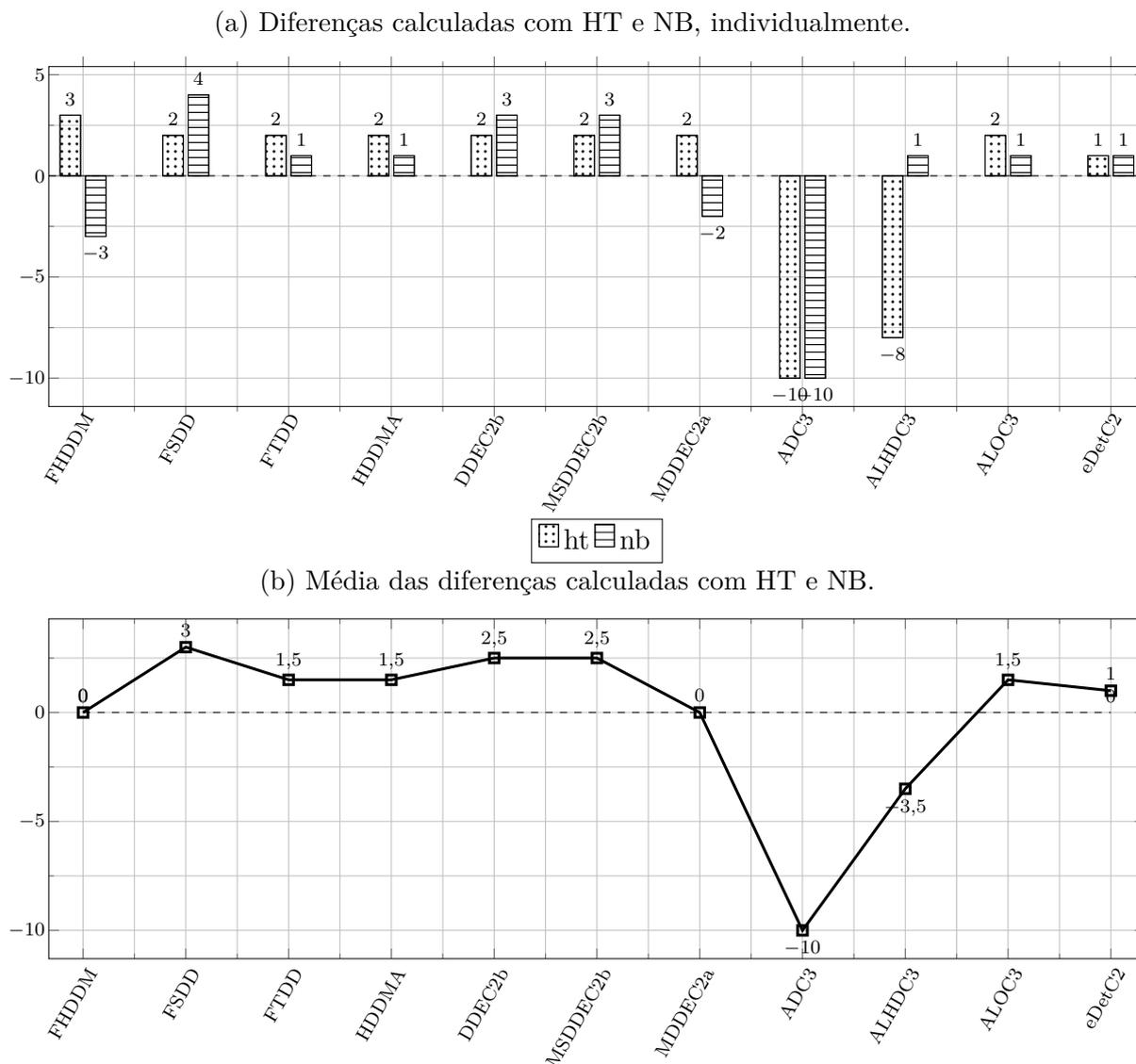


Fonte: O AUTOR (2020).

no desempenho acumulado, seguidos pelo método FHDDM com 3,5. Vale ressaltar que todos os métodos de comitês avaliados no apêndice C obtiveram rank com valor negativo (abaixo de zero). Outro ponto importante é que todos os métodos propostos obtiveram ranks com valores positivos.

As Tabelas 6 e 7 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos ALOC₂ e eDetC₂ obtiveram os melhores resultados com o HT (4, 39) e NB (4, 67), respectivamente. O mapa de calor mostra todas as linhas em vermelho nos métodos AD e ALHD com o HT e AD com o NB, implicando nos piores resultados.

Figura 6 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise.



Fonte: O AUTOR (2020).

A Figura 6 (a) traz o resumo dos resultados com HT e NB apresentados no apêndice D nas Tabelas 82 e 83. Em termos das hipóteses testadas, o método FHDDM com o HT forneceu o melhor desempenho com 3 vitórias. Tratando-se do NB, o método FSDD obteve 4 vitórias.

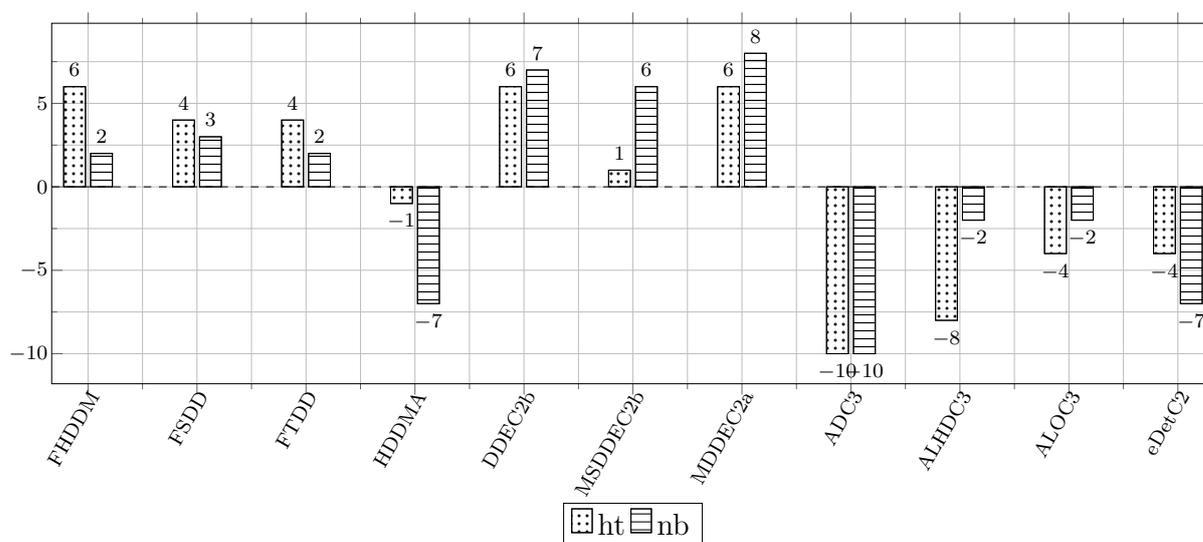
Atende-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês MDDEC_{1a} e MDDEC_{2a}, ambos com rank de avaliação acima de zero, mostrados na Figura 6 (b). Apesar disso, vale a pena destacar que o desempenho deles foi bom na maioria das situações com o HT e com o NB. A Figura 6 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que o FSDD obteve a maior média (3) no desempenho acumulado,

seguido pelo métodos DDE.C₂b e MDDEC₂a com 2.

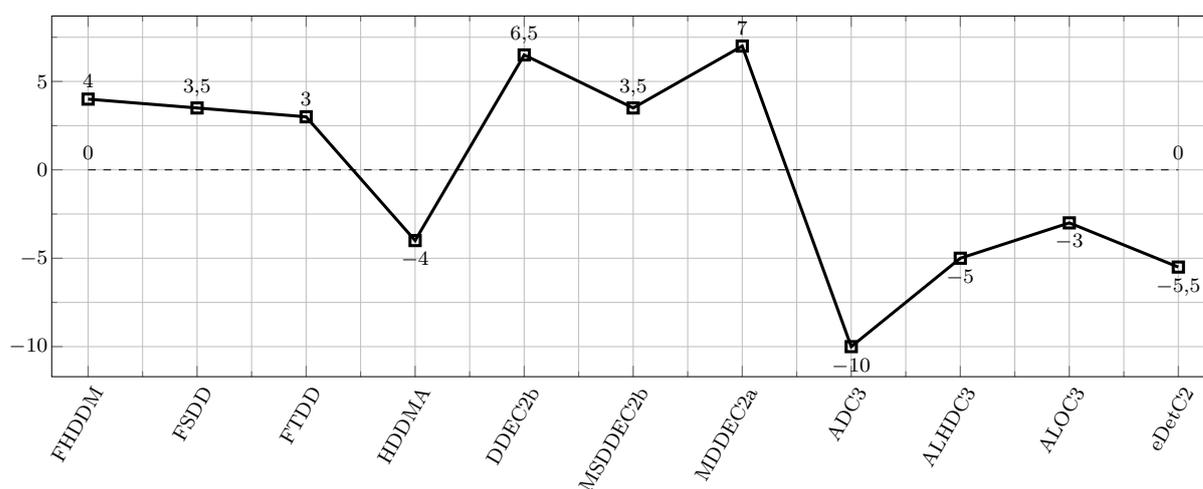
As Tabelas 8 e 9 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o método DDE.C₂a obteve os melhores resultados com HT (3, 28) e com NB foi o método MDDEC₂a com 2, 75.

Figura 7 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



Fonte: O AUTOR (2020).

A Figura 7 (a) traz o resumo dos resultados com HT e NB apresentados no apêndice D nas Tabelas 84 e 85. Em termos das hipóteses testadas, os métodos FHDDM, DDE.C₂b e MDDEC₂a com o HT obtiveram 6 vitórias. Já com o NB o MDDEC₂a obteve 8 vitórias. Em termos estatísticos, os métodos FHDDM, FSDD, FTDD e

MDDEC_{2a} foram iguais com o HT e com o NB. Já os métodos DDE.C_{2b} e MSDDE.C_{2b} foram iguais estatisticamente ao MDDEC_{2b}. Os métodos com os maiores números de derrotas usando HT foram ADC₃, ALHDC₃ e eDetC₂. Essa ineficiência mantém-se com NB, nos métodos ADC₃ e eDetC₂, além deles, aparece o HDDMA. Logo, tratando-se da ineficiência, eles foram em termos absolutos e em termos estatísticos os piores métodos.

Nota-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos métodos FHDDM, FSDD, FTDD, DDE.C_{2b}, MSDDE.C_{2b} e MDDEC_{2a}, todos com avaliações acima de zero, mostrados na Figura 7 (b). Apesar disso, vale a pena destacar que o desempenho dos comitês MDDEC_{2a} foi sólido na maioria das situações e líderes em vitórias com HT e NB, respectivamente. A Figura 7 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que o comitê MDDEC_{2a} obteve a maior média no desempenho acumulado com 7, seguido pelo método DDE.C_{2b} com 6,5. É importante destacar que novamente os métodos de comitês do apêndice C ficaram com média negativa em número de vitórias, ou seja, foram inferiores mais vezes que superiores estatisticamente.

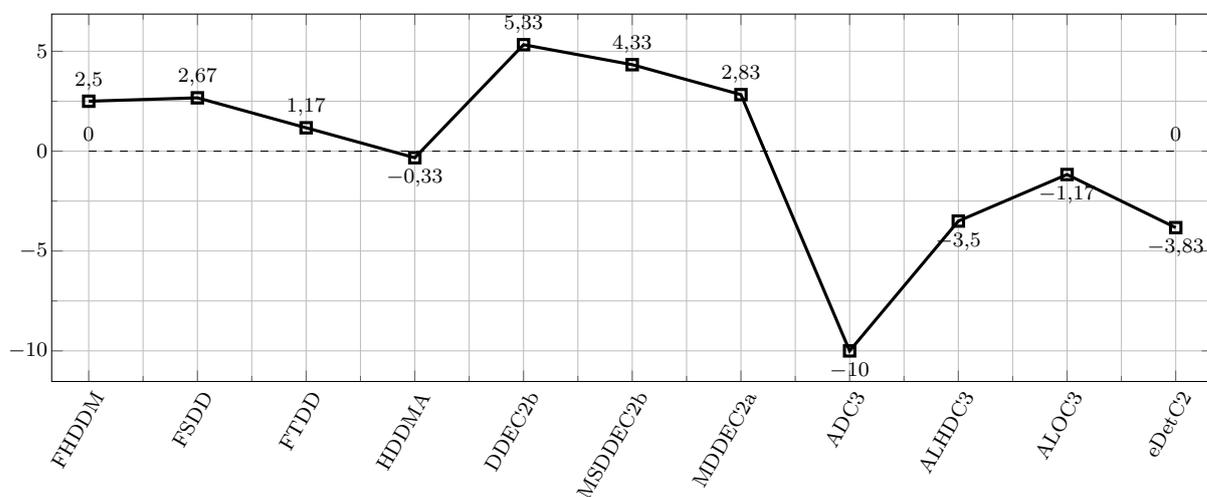
A Figura 8 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 8 (a) é feito o rank da média das diferenças. A Figura 8 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 8 (a) ou no da Figura 8 (b), o detector de maior valor absoluto foi o DDE.C_{2b} com médias 5,83 e 9,58, respectivamente. Vale ressaltar que por meio do rank da média (Figura 8 (a)) é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários.

O DDE.C_{2b} foi bem na maioria das avaliações. Apesar disso, ele foi inferior estatisticamente aos métodos FHDDM, FSDD, FTDD e MDDEC_{2a} no critério MCC com o HT. Entretanto, ele foi o melhor na classificação ordenada por cálculo das médias dos ranks e por cálculo de média dos postos. Vale ressaltar que os três métodos propostos obtiveram os melhores resultados, na qual DDE.C_{2b} com 5,83, MSDDE.C_{2b} com 4,33 e MDDEC_{2a} com 2,83. O melhor método de única análise, o FHDDM, obteve 2,5, menos da metade da média obtida pelo DDE.C_{2b}.

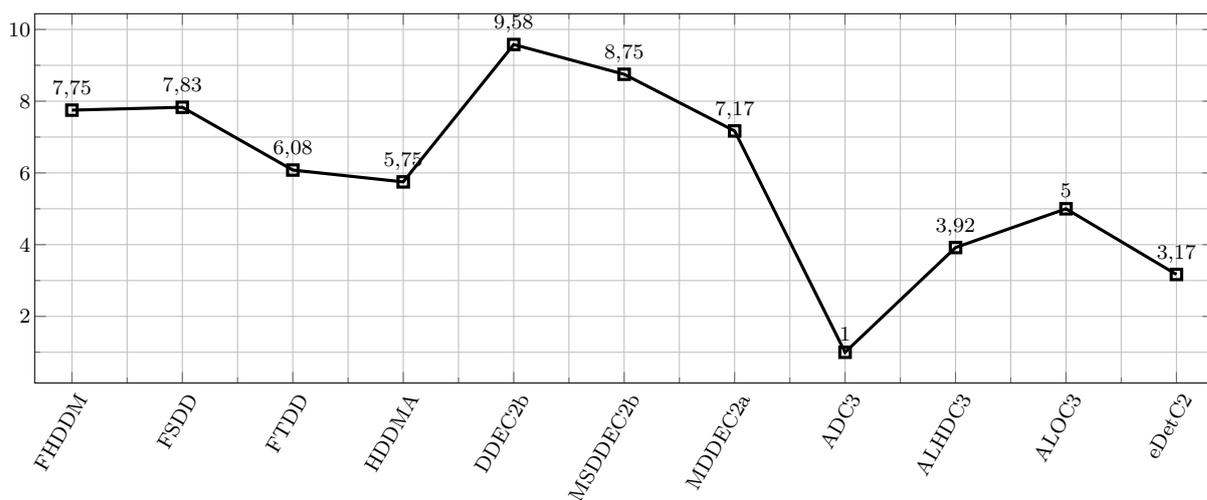
Finalmente, a Figura 9 apresenta as vitórias obtidas pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e dois classificadores (HT e NB). A leitura do gráfico é feita por método individualmente e tem por objetivo avaliar os pontos fracos e fortes de cada método nas três medidas de desempenho. Nas Figura 9 (a) e Figura 9 (b), são mostradas as vitórias com HT e NB, respectivamente. É possível observar que o detector AD não obteve vitórias em todos os critérios avaliados em ambos os classificadores. Além disso, também se observa que o MDDE alcançou um percentual muito maior (mais que 70%) em vitórias com MCC no NB, quando

Figura 8 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com os métodos propostos de múltipla análise.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.

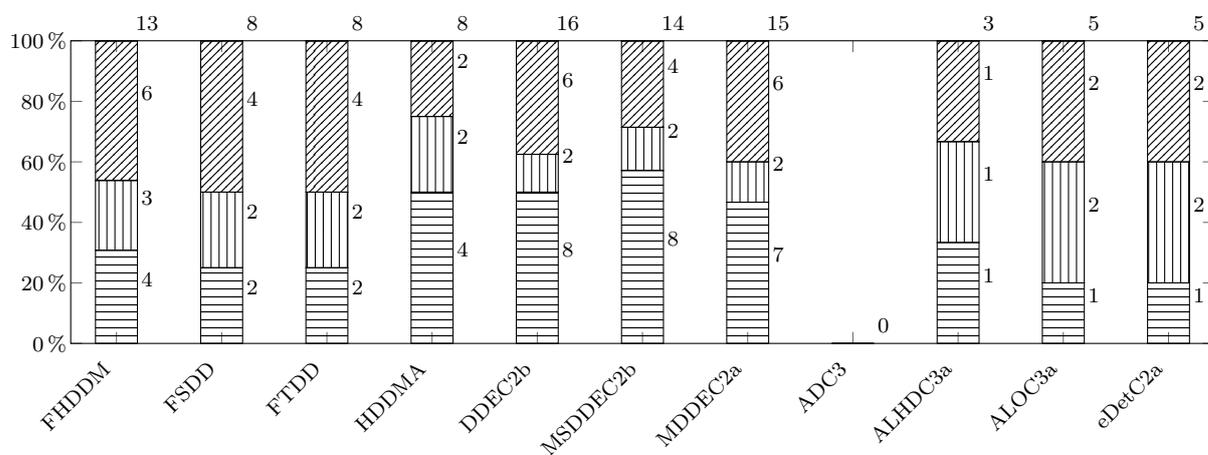


Fonte: O AUTOR (2020).

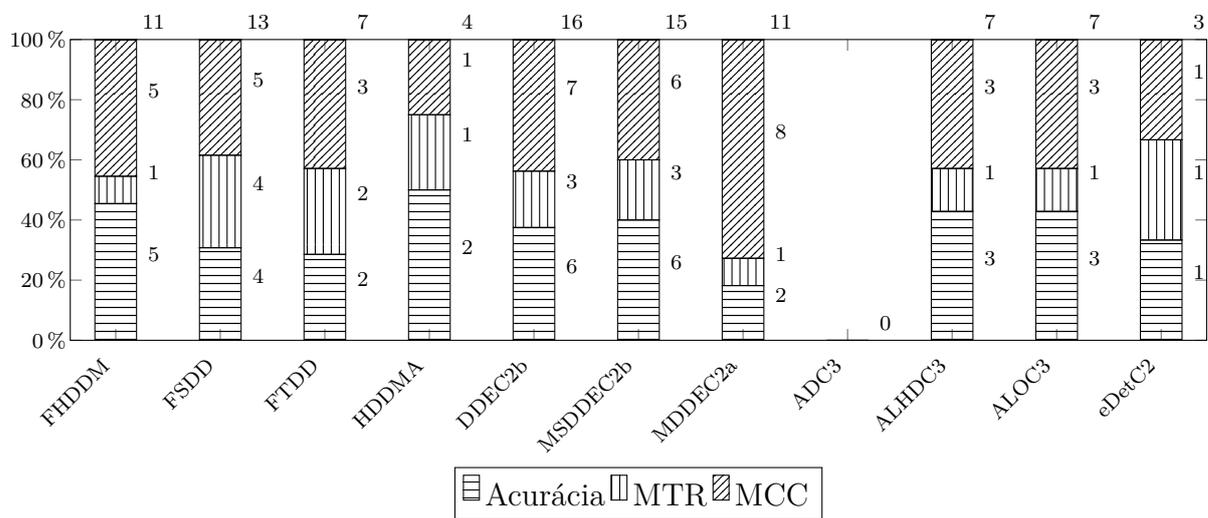
comparado em relação à acurácia e MTR. Note-se que os percentuais em termos de MTR também foram os menores quando comparado aos de acurácia e MCC entre os detectores e classificadores. Por fim, o DDE mostrou-se eficiente com HT e NB, sendo o líder no acumulado em vitórias.

Figura 9 – Resumo das vitórias calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com os métodos propostos de múltipla análise.

(a) Vitórias calculadas usando HT.



(b) Vitórias calculadas usando NB.



Acurácia
 MTR
 MCC

Fonte: O AUTOR (2020).

Tabela 4 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|--------------|--------------|--------------|
| A-20k-Agr | 86,92 | 86,44 | 86,20 | 86,67 | 86,93 | 86,92 | 87,05 | 74,72 | 86,96 | 86,96 | 82,04 |
| A-20k-LED | 69,98 | 68,48 | 68,54 | 70,02 | 69,94 | 70,09 | 70,16 | 63,31 | 68,67 | 68,67 | 70,54 |
| A-20k-Mix | 90,75 | 90,82 | 90,82 | 90,49 | 90,76 | 90,78 | 90,82 | 61,33 | 90,76 | 90,76 | 83,41 |
| A-20k-RBF | 74,83 | 74,94 | 74,81 | 74,79 | 74,96 | 74,90 | 74,91 | 72,45 | 74,84 | 74,84 | 73,12 |
| A-20k-Sine | 89,73 | 89,76 | 89,76 | 89,67 | 89,75 | 89,76 | 89,79 | 62,40 | 89,73 | 89,73 | 79,43 |
| A-20k-Wav | 79,32 | 79,38 | 79,37 | 79,29 | 79,31 | 79,29 | 79,32 | 77,67 | 79,21 | 79,21 | 78,53 |
| A-130k-Agr | 91,50 | 92,03 | 92,02 | 91,74 | 92,03 | 92,00 | 92,01 | 77,81 | 91,47 | 91,47 | 90,04 |
| A-130k-LED | 71,45 | 70,74 | 71,06 | 71,47 | 71,55 | 71,72 | 71,72 | 66,88 | 70,00 | 70,00 | 72,09 |
| A-130k-Mix | 92,60 | 92,58 | 92,58 | 92,57 | 92,59 | 92,60 | 92,61 | 61,38 | 92,58 | 92,58 | 70,84 |
| A-130k-RBF | 77,99 | 79,59 | 79,61 | 79,27 | 79,53 | 79,34 | 79,50 | 77,33 | 78,01 | 78,01 | 78,61 |
| A-130k-Sine | 93,34 | 93,34 | 93,34 | 93,28 | 93,34 | 93,33 | 93,35 | 64,75 | 93,33 | 93,33 | 81,83 |
| A-130k-Wav | 79,62 | 79,60 | 79,59 | 79,57 | 79,70 | 79,76 | 79,83 | 79,08 | 79,70 | 79,70 | 79,31 |
| A-890k-Agr | 95,16 | 95,26 | 95,26 | 95,14 | 95,26 | 95,26 | 95,26 | 78,86 | 95,15 | 95,15 | 94,97 |
| A-890k-LED | 71,71 | 71,25 | 71,70 | 71,71 | 71,87 | 71,97 | 71,97 | 68,44 | 70,35 | 70,35 | 72,44 |
| A-890k-Mix | 94,97 | 94,95 | 94,95 | 94,95 | 94,97 | 94,97 | 94,98 | 61,30 | 94,96 | 94,96 | 74,60 |
| A-890k-RBF | 83,55 | 85,09 | 85,16 | 84,37 | 85,04 | 84,97 | 84,98 | 82,45 | 83,47 | 83,47 | 84,43 |
| A-890k-Sine | 96,97 | 96,96 | 96,96 | 96,91 | 96,97 | 96,99 | 96,97 | 63,84 | 96,97 | 96,97 | 84,83 |
| A-890k-Wav | 81,22 | 81,97 | 82,17 | 81,76 | 81,67 | 81,62 | 81,73 | 81,54 | 80,85 | 80,85 | 81,58 |
| G-20k-Agr | 85,17 | 84,32 | 83,92 | 85,30 | 84,96 | 85,07 | 85,09 | 73,83 | 85,23 | 85,23 | 81,09 |
| G-20k-LED | 69,65 | 68,17 | 68,49 | 69,54 | 69,48 | 69,66 | 69,63 | 63,20 | 68,30 | 68,30 | 69,86 |
| G-20k-Mix | 87,77 | 87,25 | 87,22 | 87,48 | 87,55 | 87,41 | 87,32 | 61,76 | 87,43 | 87,43 | 87,01 |
| G-20k-RBF | 73,91 | 74,04 | 73,78 | 74,13 | 74,10 | 74,03 | 74,08 | 72,14 | 73,83 | 73,83 | 72,95 |
| G-20k-Sine | 86,99 | 86,53 | 86,47 | 86,74 | 86,76 | 86,72 | 86,69 | 63,75 | 86,70 | 86,70 | 85,67 |
| G-20k-Wav | 78,79 | 78,60 | 78,40 | 78,66 | 78,72 | 78,77 | 78,73 | 77,58 | 78,77 | 78,77 | 78,39 |
| G-130k-Agr | 91,19 | 91,74 | 91,72 | 91,51 | 91,74 | 91,71 | 91,72 | 78,81 | 91,19 | 91,19 | 90,74 |
| G-130k-LED | 71,39 | 70,76 | 70,98 | 71,36 | 71,48 | 71,65 | 71,59 | 66,87 | 69,87 | 69,87 | 72,05 |
| G-130k-Mix | 92,14 | 92,03 | 92,03 | 92,06 | 92,09 | 92,08 | 92,07 | 62,09 | 92,04 | 92,04 | 92,06 |
| G-130k-RBF | 77,68 | 79,38 | 79,41 | 79,15 | 79,27 | 79,19 | 79,22 | 77,33 | 77,79 | 77,79 | 78,81 |
| G-130k-Sine | 92,84 | 92,77 | 92,77 | 92,77 | 92,81 | 92,80 | 92,81 | 66,51 | 92,78 | 92,78 | 92,74 |
| G-130k-Wav | 79,68 | 79,63 | 79,84 | 79,59 | 79,72 | 79,75 | 79,72 | 79,02 | 79,71 | 79,71 | 79,33 |
| G-890k-Agr | 95,10 | 95,23 | 95,23 | 95,10 | 95,22 | 95,22 | 95,23 | 79,01 | 95,11 | 95,11 | 95,13 |
| G-890k-LED | 71,70 | 71,23 | 71,69 | 71,72 | 71,87 | 71,96 | 71,95 | 68,70 | 70,34 | 70,34 | 72,44 |
| G-890k-Mix | 94,90 | 94,85 | 94,86 | 94,88 | 94,88 | 94,88 | 94,87 | 62,89 | 94,85 | 94,85 | 94,88 |
| G-890k-RBF | 83,23 | 85,07 | 85,09 | 84,30 | 84,99 | 84,81 | 84,90 | 82,73 | 83,07 | 83,07 | 84,49 |
| G-890k-Sine | 96,89 | 96,84 | 96,84 | 96,85 | 96,87 | 96,89 | 96,86 | 66,77 | 96,86 | 96,86 | 96,86 |
| G-890k-Wav | 81,04 | 81,64 | 82,00 | 81,66 | 81,68 | 81,57 | 81,69 | 81,52 | 80,86 | 80,86 | 81,56 |
| Airlines | 65,73 | 65,71 | 65,77 | 66,00 | 65,93 | 65,98 | 65,94 | 65,44 | 65,67 | 65,67 | 66,46 |
| Connect4 | 75,21 | 74,88 | 74,57 | 75,14 | 75,03 | 75,09 | 75,05 | 74,00 | 75,10 | 75,10 | 74,91 |
| Outdoor | 60,09 | 58,46 | 59,21 | 59,96 | 60,31 | 60,18 | 59,96 | 58,42 | 61,14 | 61,14 | 61,27 |
| Rialto | 41,75 | 37,10 | 33,45 | 44,24 | 43,03 | 38,65 | 38,21 | 31,10 | 39,98 | 39,98 | 38,81 |
| SpamData | 92,57 | 92,18 | 92,18 | 92,76 | 92,17 | 92,19 | 92,42 | 91,57 | 92,16 | 92,16 | 92,32 |
| WineRed | 52,07 | 52,71 | 53,17 | 52,82 | 52,71 | 52,55 | 51,16 | 50,96 | 52,24 | 52,24 | 51,33 |
| WineWhite | 46,09 | 44,44 | 44,44 | 45,15 | 44,41 | 44,42 | 44,42 | 45,13 | 47,14 | 47,14 | 44,80 |
| Vitórias | 4 | 2 | 2 | 4 | 8 | 8 | 7 | 0 | 1 | 1 | 1 |
| Derrotas | 2 | 3 | 3 | 3 | 0 | 0 | 0 | 10 | 5 | 5 | 7 |
| Média Postos | 5,10 | 5,98 | 5,79 | 5,74 | 4,00 | 4,17 | 3,87 | 10,72 | 6,80 | 6,80 | 7,01 |

Fonte: O AUTOR (2020).

Tabela 5 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|--------------|--------------|--------------|
| A-20k-Agr | 86,95 | 86,83 | 85,71 | 86,86 | 87,03 | 87,02 | 86,72 | 68,42 | 86,99 | 86,99 | 79,51 |
| A-20k-LED | 71,04 | 67,97 | 66,10 | 71,40 | 70,59 | 70,77 | 67,38 | 60,11 | 69,74 | 69,74 | 69,33 |
| A-20k-Mix | 91,11 | 91,19 | 91,19 | 91,10 | 91,14 | 91,15 | 91,17 | 57,99 | 91,13 | 91,13 | 75,18 |
| A-20k-RBF | 74,88 | 74,89 | 74,83 | 74,77 | 74,86 | 74,81 | 74,83 | 63,28 | 74,86 | 74,86 | 70,55 |
| A-20k-Sine | 89,15 | 89,18 | 89,18 | 89,04 | 89,16 | 89,17 | 89,18 | 59,24 | 89,15 | 89,15 | 70,47 |
| A-20k-Wav | 79,70 | 79,69 | 79,55 | 79,46 | 79,67 | 79,65 | 79,67 | 75,73 | 79,70 | 79,70 | 78,23 |
| A-130k-Agr | 88,71 | 88,91 | 88,90 | 88,75 | 88,90 | 88,89 | 88,91 | 69,12 | 88,72 | 88,72 | 86,90 |
| A-130k-LED | 72,63 | 72,62 | 70,99 | 73,44 | 73,05 | 73,25 | 71,45 | 60,33 | 71,94 | 71,94 | 73,03 |
| A-130k-Mix | 91,93 | 91,94 | 91,94 | 91,84 | 91,93 | 91,93 | 91,94 | 58,34 | 91,93 | 91,93 | 65,30 |
| A-130k-RBF | 75,26 | 75,31 | 75,28 | 75,25 | 75,32 | 75,28 | 75,04 | 63,61 | 75,25 | 75,25 | 74,36 |
| A-130k-Sine | 89,48 | 89,49 | 89,49 | 89,37 | 89,48 | 89,49 | 89,49 | 59,88 | 89,48 | 89,48 | 62,54 |
| A-130k-Wav | 80,33 | 80,33 | 80,30 | 80,22 | 80,32 | 80,32 | 80,34 | 75,48 | 80,32 | 80,32 | 79,77 |
| A-890k-Agr | 89,10 | 89,36 | 89,37 | 89,30 | 89,34 | 89,33 | 89,36 | 69,34 | 89,10 | 89,10 | 88,99 |
| A-890k-LED | 73,02 | 73,82 | 73,70 | 73,87 | 73,81 | 73,80 | 73,43 | 60,61 | 72,49 | 72,49 | 73,78 |
| A-890k-Mix | 92,07 | 92,08 | 92,08 | 92,03 | 92,07 | 92,07 | 92,07 | 58,41 | 92,07 | 92,07 | 60,27 |
| A-890k-RBF | 75,36 | 75,46 | 75,46 | 75,43 | 75,44 | 75,45 | 75,41 | 63,20 | 75,36 | 75,36 | 75,32 |
| A-890k-Sine | 89,50 | 89,50 | 89,50 | 89,45 | 89,50 | 89,50 | 89,50 | 59,47 | 89,50 | 89,50 | 75,73 |
| A-890k-Wav | 80,38 | 80,40 | 80,40 | 80,35 | 80,40 | 80,39 | 80,40 | 76,09 | 80,37 | 80,37 | 80,30 |
| G-20k-Agr | 84,07 | 82,95 | 80,48 | 84,79 | 84,12 | 84,72 | 80,44 | 67,93 | 84,23 | 84,23 | 78,04 |
| G-20k-LED | 70,40 | 67,11 | 64,83 | 71,07 | 69,42 | 69,82 | 66,02 | 60,09 | 69,26 | 69,26 | 69,00 |
| G-20k-Mix | 88,18 | 87,67 | 87,65 | 87,81 | 87,91 | 87,76 | 87,82 | 58,29 | 87,83 | 87,83 | 87,63 |
| G-20k-RBF | 73,79 | 73,75 | 73,16 | 73,91 | 73,76 | 73,61 | 72,66 | 63,46 | 73,78 | 73,78 | 71,21 |
| G-20k-Sine | 87,11 | 86,55 | 86,56 | 86,96 | 87,00 | 86,91 | 86,89 | 59,46 | 86,75 | 86,75 | 85,66 |
| G-20k-Wav | 78,95 | 78,93 | 78,92 | 78,97 | 78,97 | 78,97 | 78,84 | 75,43 | 78,94 | 78,94 | 77,78 |
| G-130k-Agr | 88,31 | 88,39 | 88,36 | 88,45 | 88,31 | 88,33 | 87,50 | 69,12 | 88,33 | 88,33 | 87,08 |
| G-130k-LED | 72,54 | 72,11 | 70,35 | 73,41 | 72,52 | 72,77 | 70,90 | 60,58 | 71,87 | 71,87 | 73,11 |
| G-130k-Mix | 91,47 | 91,40 | 91,39 | 91,42 | 91,43 | 91,42 | 91,42 | 58,36 | 91,40 | 91,40 | 91,40 |
| G-130k-RBF | 75,06 | 75,07 | 75,06 | 75,11 | 75,07 | 75,05 | 74,77 | 63,68 | 75,05 | 75,05 | 74,70 |
| G-130k-Sine | 89,31 | 89,19 | 89,18 | 89,28 | 89,31 | 89,31 | 89,31 | 59,04 | 89,19 | 89,19 | 89,10 |
| G-130k-Wav | 80,11 | 80,08 | 80,02 | 80,09 | 80,08 | 80,08 | 80,05 | 75,79 | 80,10 | 80,10 | 79,76 |
| G-890k-Agr | 89,02 | 89,26 | 89,27 | 89,23 | 89,26 | 89,22 | 88,51 | 69,30 | 89,02 | 89,02 | 89,05 |
| G-890k-LED | 73,01 | 73,77 | 73,63 | 73,87 | 73,53 | 73,73 | 72,78 | 60,94 | 72,48 | 72,48 | 73,78 |
| G-890k-Mix | 92,00 | 92,00 | 91,99 | 91,99 | 92,00 | 92,00 | 92,00 | 58,41 | 91,99 | 91,99 | 92,00 |
| G-890k-RBF | 75,34 | 75,42 | 75,42 | 75,42 | 75,41 | 75,42 | 75,18 | 64,10 | 75,33 | 75,33 | 75,36 |
| G-890k-Sine | 89,50 | 89,47 | 89,48 | 89,45 | 89,30 | 89,50 | 89,51 | 59,53 | 89,47 | 89,47 | 89,46 |
| G-890k-Wav | 80,35 | 80,35 | 80,35 | 80,34 | 80,34 | 80,34 | 80,34 | 76,22 | 80,34 | 80,34 | 80,32 |
| Airlines | 65,82 | 66,91 | 66,76 | 67,23 | 66,96 | 66,97 | 67,05 | 65,41 | 65,66 | 65,66 | 66,81 |
| Connect4 | 75,23 | 74,65 | 74,23 | 74,95 | 75,11 | 75,02 | 74,57 | 70,88 | 75,07 | 75,07 | 74,47 |
| Outdoor | 60,59 | 60,23 | 59,80 | 60,11 | 60,63 | 60,32 | 60,71 | 58,53 | 61,59 | 61,59 | 61,23 |
| Rialto | 41,24 | 28,12 | 27,52 | 43,30 | 34,67 | 30,13 | 24,45 | 21,59 | 36,84 | 36,84 | 35,86 |
| SpamData | 91,51 | 91,35 | 91,35 | 90,96 | 91,22 | 91,58 | 91,32 | 90,62 | 91,41 | 91,41 | 91,36 |
| WineRed | 47,46 | 47,85 | 47,70 | 47,63 | 47,51 | 47,35 | 47,52 | 47,70 | 47,82 | 47,82 | 47,76 |
| WineWhite | 46,53 | 45,48 | 44,88 | 43,16 | 45,62 | 45,90 | 46,64 | 42,82 | 46,67 | 46,67 | 46,72 |
| Vitórias | 5 | 4 | 2 | 2 | 6 | 6 | 2 | 0 | 3 | 3 | 1 |
| Derrotas | 0 | 0 | 3 | 0 | 0 | 0 | 6 | 10 | 3 | 3 | 9 |
| Média Postos | 4,66 | 4,50 | 6,05 | 5,31 | 4,52 | 4,56 | 5,99 | 10,87 | 5,80 | 5,80 | 7,93 |

Fonte: O AUTOR (2020).

Tabela 6 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|--------------|-----------------|-----------------|-----------------|---------------|----------------|----------|-------|---------|---------------|----------------|
| A-20k-Agr | 104,44 | 64,39 | 49,07 | 63,50 | 104,95 | 109,26 | 112,53 | 0,00 | 0,26 | 123,84 | 93,69 |
| A-20k-LED | 7,63 | 1,32 | 1,66 | 2,74 | 2,51 | 3,35 | 2,53 | 0,00 | 0,06 | 4,71 | 4,24 |
| A-20k-Mix | 310,41 | 340,77 | 340,77 | 257,89 | 317,56 | 317,72 | 320,09 | 0,00 | 0,17 | 334,99 | 198,63 |
| A-20k-RBF | 32,20 | 7,39 | 4,95 | 14,85 | 16,30 | 16,88 | 16,48 | 0,00 | 0,17 | 31,75 | 30,88 |
| A-20k-Sine | 283,61 | 285,17 | 285,17 | 252,81 | 286,89 | 287,41 | 287,25 | 0,00 | 0,22 | 305,81 | 212,73 |
| A-20k-Wav | 1,47 | 1,08 | 0,84 | 1,20 | 0,80 | 0,84 | 0,93 | 0,01 | 0,09 | 1,77 | 0,99 |
| A-130k-Agr | 732,18 | 1069,78 | 1011,47 | 702,40 | 995,02 | 1003,12 | 1011,64 | 0,00 | 0,20 | 965,99 | 366,78 |
| A-130k-LED | 23,39 | 1,45 | 2,12 | 16,49 | 7,45 | 10,73 | 10,34 | 0,00 | 0,10 | 9,68 | 8,92 |
| A-130k-Mix | 1993,70 | 1970,80 | 1970,80 | 2497,95 | 2067,37 | 2081,44 | 2100,45 | 0,00 | 114,80 | 1920,33 | 1129,24 |
| A-130k-RBF | 207,39 | 286,43 | 243,09 | 218,18 | 351,52 | 315,34 | 309,98 | 0,00 | 0,16 | 225,27 | 219,24 |
| A-130k-Sine | 1772,46 | 2011,33 | 2011,33 | 2568,86 | 1989,87 | 1992,19 | 1991,98 | 0,01 | 0,12 | 2022,41 | 960,66 |
| A-130k-Wav | 63,78 | 23,80 | 8,42 | 21,85 | 42,51 | 41,07 | 45,90 | 0,01 | 0,20 | 43,54 | 88,43 |
| A-890k-Agr | 5475,15 | 9069,74 | 9114,79 | 2518,23 | 6654,72 | 6568,56 | 6661,40 | 0,01 | 0,51 | 7759,54 | 2523,80 |
| A-890k-LED | 25,54 | 2,34 | 1,63 | 15,16 | 10,79 | 11,53 | 7,52 | 0,00 | 0,31 | 9,94 | 9,44 |
| A-890k-Mix | 13538,84 | 16451,68 | 16451,68 | 16981,12 | 14973,11 | 15188,82 | 15368,69 | 0,01 | 1587,13 | 15870,89 | 8163,69 |
| A-890k-RBF | 1310,04 | 2075,84 | 2412,69 | 1099,40 | 2092,24 | 1798,96 | 2011,13 | 0,01 | 0,28 | 1400,76 | 702,05 |
| A-890k-Sine | 11830,04 | 16724,42 | 16724,42 | 13992,72 | 14485,30 | 14406,21 | 14445,17 | 0,00 | 0,19 | 16288,62 | 8918,58 |
| A-890k-Wav | 218,15 | 449,29 | 298,90 | 556,74 | 159,77 | 173,85 | 118,63 | 0,00 | 0,25 | 47,74 | 128,59 |
| G-20k-Agr | 15,47 | 4,08 | 1,64 | 7,43 | 8,81 | 10,11 | 11,32 | 0,00 | 0,06 | 9,79 | 7,72 |
| G-20k-LED | 2,29 | 1,27 | 1,53 | 2,05 | 2,15 | 2,17 | 2,17 | 0,00 | 0,02 | 1,64 | 1,60 |
| G-20k-Mix | 0,95 | 5,49 | 4,98 | 2,68 | 0,84 | 0,85 | 0,86 | 0,00 | 0,10 | 6,24 | 12,27 |
| G-20k-RBF | 6,86 | 1,55 | 1,20 | 7,24 | 4,96 | 4,76 | 5,78 | 0,00 | 0,06 | 6,78 | 6,63 |
| G-20k-Sine | 1,32 | 2,17 | 2,08 | 1,20 | 1,05 | 1,08 | 0,98 | 0,00 | 0,20 | 3,24 | 14,46 |
| G-20k-Wav | 0,64 | 0,39 | 0,28 | 0,47 | 0,42 | 0,52 | 0,49 | 0,00 | 0,04 | 0,65 | 0,54 |
| G-130k-Agr | 129,63 | 107,01 | 96,21 | 99,82 | 180,39 | 171,53 | 160,33 | 0,00 | 0,26 | 89,49 | 63,21 |
| G-130k-LED | 8,38 | 1,73 | 1,71 | 4,93 | 3,20 | 3,99 | 3,69 | 0,00 | 0,05 | 2,47 | 3,61 |
| G-130k-Mix | 0,87 | 32,21 | 29,52 | 45,98 | 0,90 | 0,98 | 1,09 | 0,00 | 0,11 | 28,06 | 92,13 |
| G-130k-RBF | 42,27 | 62,54 | 48,59 | 86,37 | 60,55 | 52,50 | 59,36 | 0,00 | 0,16 | 48,00 | 40,13 |
| G-130k-Sine | 1,50 | 20,95 | 8,42 | 0,87 | 1,00 | 1,11 | 1,19 | 0,00 | 0,16 | 21,29 | 122,01 |
| G-130k-Wav | 3,40 | 6,20 | 1,04 | 9,29 | 4,09 | 3,63 | 3,64 | 0,00 | 0,07 | 3,35 | 7,08 |
| G-890k-Agr | 1091,11 | 661,56 | 661,55 | 538,99 | 1518,90 | 1699,71 | 1455,29 | 0,01 | 0,36 | 799,73 | 536,85 |
| G-890k-LED | 6,91 | 1,22 | 1,65 | 5,38 | 7,32 | 4,16 | 5,22 | 0,00 | 0,32 | 3,13 | 3,11 |
| G-890k-Mix | 173,78 | 294,47 | 308,53 | 0,95 | 0,80 | 0,93 | 0,90 | 0,00 | 0,18 | 383,43 | 1593,44 |
| G-890k-RBF | 241,91 | 526,61 | 573,54 | 289,71 | 430,97 | 383,51 | 467,81 | 0,01 | 0,34 | 246,21 | 213,02 |
| G-890k-Sine | 154,03 | 14,31 | 14,32 | 0,85 | 0,86 | 0,92 | 0,91 | 0,00 | 0,24 | 46,50 | 1272,62 |
| G-890k-Wav | 17,08 | 44,76 | 47,32 | 145,44 | 59,01 | 32,86 | 59,91 | 0,00 | 0,34 | 17,03 | 74,96 |
| Vitórias | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 2 | 2 |
| Derrotas | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 9 | 0 | 1 |
| Média Postos | 4,72 | 5,00 | 5,65 | 4,97 | 5,28 | 4,78 | 4,60 | 11,00 | 10,00 | 4,39 | 5,61 |

Fonte: O AUTOR (2020).

Tabela 7 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDECC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|--------------|----------------|----------------|---------------|---------------|---------------|---------------|-------|-----------------|-----------------|-----------------|
| A-20k-Agr | 116,51 | 94,07 | 66,48 | 87,95 | 118,32 | 125,47 | 86,87 | 0,03 | 126,37 | 126,37 | 114,84 |
| A-20k-LED | 44,97 | 0,81 | 0,45 | 19,59 | 15,77 | 16,00 | 0,90 | 0,01 | 30,52 | 30,52 | 28,85 |
| A-20k-Mix | 306,08 | 345,14 | 345,14 | 459,62 | 324,63 | 325,55 | 306,08 | 0,00 | 341,21 | 341,21 | 487,89 |
| A-20k-RBF | 25,85 | 10,34 | 3,62 | 20,44 | 14,65 | 15,17 | 9,93 | 0,02 | 24,68 | 24,68 | 26,86 |
| A-20k-Sine | 284,58 | 293,51 | 288,32 | 221,29 | 288,20 | 288,26 | 265,71 | 0,07 | 311,75 | 311,75 | 321,49 |
| A-20k-Wav | 51,84 | 34,64 | 22,15 | 27,17 | 40,73 | 43,38 | 31,51 | 0,06 | 44,00 | 44,00 | 4,43 |
| A-130k-Agr | 531,93 | 721,14 | 594,73 | 289,91 | 673,58 | 675,38 | 617,18 | 0,02 | 656,66 | 656,66 | 336,73 |
| A-130k-LED | 59,68 | 26,87 | 4,86 | 87,89 | 100,88 | 101,39 | 34,57 | 0,03 | 25,69 | 25,69 | 64,09 |
| A-130k-Mix | 2037,77 | 2277,70 | 2277,70 | 1375,12 | 2099,09 | 2102,18 | 2037,77 | 0,00 | 2234,19 | 2234,19 | 2347,58 |
| A-130k-RBF | 89,45 | 136,42 | 56,55 | 89,95 | 166,44 | 151,33 | 130,90 | 0,09 | 87,01 | 87,01 | 92,68 |
| A-130k-Sine | 1894,08 | 1987,90 | 1953,63 | 642,88 | 1811,29 | 1814,46 | 1768,44 | 0,09 | 2112,52 | 2112,52 | 1975,02 |
| A-130k-Wav | 447,25 | 496,92 | 367,31 | 199,48 | 463,13 | 455,83 | 470,34 | 0,03 | 413,18 | 413,18 | 107,76 |
| A-890k-Agr | 4038,41 | 4630,12 | 4142,22 | 765,94 | 4024,34 | 3886,66 | 4114,44 | 0,02 | 4096,23 | 4096,23 | 1495,59 |
| A-890k-LED | 76,17 | 218,26 | 112,85 | 339,96 | 365,71 | 302,68 | 198,39 | 0,05 | 34,23 | 34,23 | 66,94 |
| A-890k-Mix | 13633,88 | 14959,61 | 14959,61 | 4215,07 | 13760,01 | 13760,01 | 13633,88 | 0,00 | 15050,81 | 15050,81 | 15795,99 |
| A-890k-RBF | 227,29 | 465,73 | 323,00 | 211,93 | 415,04 | 418,73 | 468,39 | 0,03 | 226,77 | 226,77 | 276,83 |
| A-890k-Sine | 11627,48 | 12724,30 | 12527,86 | 1368,98 | 11429,14 | 11344,26 | 11424,12 | 0,06 | 13016,89 | 13016,89 | 6401,19 |
| A-890k-Wav | 1070,72 | 1859,05 | 2051,98 | 695,58 | 1841,17 | 1723,75 | 1942,70 | 0,09 | 1028,78 | 1028,78 | 1137,73 |
| G-20k-Agr | 13,21 | 5,48 | 3,87 | 13,13 | 10,37 | 13,50 | 11,92 | 0,01 | 6,39 | 6,39 | 7,78 |
| G-20k-LED | 6,21 | 0,65 | 0,28 | 8,91 | 2,82 | 4,83 | 0,57 | 0,01 | 3,05 | 3,05 | 8,43 |
| G-20k-Mix | 1,06 | 3,53 | 3,48 | 0,88 | 0,79 | 0,83 | 1,07 | 0,01 | 3,42 | 3,42 | 19,60 |
| G-20k-RBF | 4,94 | 3,58 | 2,68 | 9,94 | 7,34 | 7,06 | 3,19 | 0,03 | 4,81 | 4,81 | 7,16 |
| G-20k-Sine | 1,45 | 2,30 | 3,72 | 0,89 | 0,89 | 0,98 | 4,01 | 0,06 | 3,24 | 3,24 | 13,97 |
| G-20k-Wav | 0,45 | 0,30 | 0,30 | 0,49 | 0,35 | 0,36 | 0,30 | 0,01 | 0,43 | 0,43 | 0,64 |
| G-130k-Agr | 76,86 | 65,85 | 57,85 | 132,99 | 142,62 | 147,32 | 126,42 | 0,02 | 45,44 | 45,44 | 40,59 |
| G-130k-LED | 13,38 | 9,89 | 1,19 | 71,71 | 33,65 | 29,13 | 7,80 | 0,05 | 6,36 | 6,36 | 19,91 |
| G-130k-Mix | 0,88 | 33,44 | 30,32 | 99,87 | 0,93 | 1,09 | 0,87 | 0,00 | 21,78 | 21,78 | 150,85 |
| G-130k-RBF | 17,44 | 24,40 | 24,94 | 51,30 | 22,57 | 26,24 | 18,67 | 0,08 | 17,75 | 17,75 | 25,37 |
| G-130k-Sine | 1,08 | 8,57 | 12,23 | 41,45 | 0,94 | 1,06 | 1,08 | 0,03 | 13,94 | 13,94 | 135,54 |
| G-130k-Wav | 64,74 | 76,30 | 48,06 | 62,82 | 66,45 | 63,07 | 59,05 | 0,09 | 72,67 | 72,67 | 23,76 |
| G-890k-Agr | 411,57 | 511,37 | 453,50 | 464,38 | 807,85 | 805,55 | 669,94 | 0,02 | 220,29 | 220,29 | 280,37 |
| G-890k-LED | 17,73 | 67,03 | 49,08 | 273,23 | 111,74 | 93,01 | 38,94 | 0,07 | 9,85 | 9,85 | 24,02 |
| G-890k-Mix | 0,81 | 223,79 | 220,70 | 1062,98 | 0,80 | 0,80 | 0,80 | 0,00 | 268,25 | 268,25 | 1527,70 |
| G-890k-RBF | 30,89 | 131,11 | 162,65 | 193,52 | 118,59 | 131,18 | 87,79 | 0,11 | 19,60 | 19,60 | 68,65 |
| G-890k-Sine | 1,39 | 12,11 | 178,08 | 409,75 | 1,14 | 1,44 | 105,30 | 0,08 | 38,22 | 38,22 | 1060,42 |
| G-890k-Wav | 124,52 | 137,51 | 403,37 | 365,24 | 226,62 | 229,37 | 254,68 | 0,11 | 83,70 | 83,70 | 150,44 |
| Vitórias | 1 | 4 | 2 | 1 | 3 | 3 | 1 | 0 | 1 | 1 | 1 |
| Derrotas | 4 | 0 | 1 | 0 | 0 | 0 | 3 | 10 | 0 | 0 | 0 |
| Média Postos | 6,17 | 4,80 | 6,10 | 5,32 | 5,42 | 4,90 | 6,31 | 11,00 | 5,67 | 5,67 | 4,67 |

Fonte: O AUTOR (2020).

Tabela 8 – Médias dos MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|-------------|-------------|-------------|-------------|-------------|----------|-------------|-------|--------|-------|--------|
| A-20k-Agr | 0,88 | 0,83 | 0,79 | 0,81 | 0,94 | 0,87 | 0,94 | 0,00 | 0,26 | 0,86 | 0,74 |
| A-20k-LED | 0,23 | 0,03 | 0,02 | 0,13 | 0,14 | 0,17 | 0,14 | 0,00 | 0,04 | 0,20 | 0,18 |
| A-20k-Mix | 1,00 | 0,95 | 0,95 | 0,97 | 1,00 | 0,95 | 1,00 | 0,00 | 0,18 | 0,95 | 0,74 |
| A-20k-RBF | 0,62 | 0,66 | 0,61 | 0,47 | 0,66 | 0,64 | 0,66 | 0,00 | 0,17 | 0,60 | 0,59 |
| A-20k-Sine | 1,00 | 0,98 | 0,98 | 1,00 | 1,00 | 0,98 | 0,99 | 0,00 | 0,17 | 0,96 | 0,79 |
| A-20k-Wav | 0,48 | 0,52 | 0,56 | 0,39 | 0,38 | 0,32 | 0,39 | 0,01 | 0,09 | 0,28 | 0,40 |
| A-130k-Agr | 0,80 | 0,94 | 0,93 | 0,87 | 0,97 | 0,82 | 0,97 | 0,00 | 0,24 | 0,80 | 0,70 |
| A-130k-LED | 0,11 | 0,01 | 0,02 | 0,07 | 0,06 | 0,08 | 0,05 | 0,00 | 0,03 | 0,09 | 0,09 |
| A-130k-Mix | 1,00 | 0,94 | 0,94 | 0,97 | 0,98 | 0,94 | 1,00 | 0,00 | 0,21 | 0,92 | 0,71 |
| A-130k-RBF | 0,45 | 0,70 | 0,69 | 0,53 | 0,73 | 0,45 | 0,70 | 0,00 | 0,19 | 0,45 | 0,48 |
| A-130k-Sine | 1,00 | 0,98 | 0,98 | 0,96 | 1,00 | 0,98 | 1,00 | 0,01 | 0,17 | 0,94 | 0,72 |
| A-130k-Wav | 0,38 | 0,44 | 0,53 | 0,19 | 0,34 | 0,28 | 0,36 | 0,01 | 0,11 | 0,21 | 0,35 |
| A-890k-Agr | 0,84 | 0,97 | 0,98 | 0,69 | 0,98 | 0,84 | 0,98 | 0,01 | 0,17 | 0,73 | 0,71 |
| A-890k-LED | 0,05 | 0,01 | 0,00 | 0,03 | 0,03 | 0,03 | 0,02 | 0,00 | 0,02 | 0,03 | 0,03 |
| A-890k-Mix | 1,00 | 0,94 | 0,94 | 0,95 | 0,98 | 0,94 | 0,99 | 0,01 | 0,26 | 0,92 | 0,70 |
| A-890k-RBF | 0,33 | 0,70 | 0,74 | 0,42 | 0,70 | 0,33 | 0,66 | 0,01 | 0,19 | 0,28 | 0,42 |
| A-890k-Sine | 1,00 | 0,98 | 0,98 | 0,97 | 1,00 | 0,98 | 1,00 | 0,00 | 0,14 | 0,95 | 0,75 |
| A-890k-Wav | 0,25 | 0,38 | 0,50 | 0,26 | 0,22 | 0,17 | 0,24 | 0,00 | 0,06 | 0,11 | 0,20 |
| Vitórias | 7 | 4 | 4 | 2 | 4 | 4 | 7 | 0 | 1 | 2 | 2 |
| Derrotas | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 10 | 9 | 6 | 6 |
| Média Postos | 3,33 | 4,92 | 5,11 | 5,67 | 3,28 | 5,81 | 3,42 | 10,98 | 9,64 | 6,92 | 6,94 |

Fonte: O AUTOR (2020).

Tabela 9 – Médias dos MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------|--------|-------|--------|
| A-20k-Agr | 0,91 | 0,87 | 0,77 | 0,84 | 0,93 | 0,97 | 0,90 | 0,05 | 0,87 | 0,87 | 0,74 |
| A-20k-LED | 0,49 | 0,10 | 0,07 | 0,17 | 0,27 | 0,28 | 0,17 | 0,00 | 0,32 | 0,32 | 0,38 |
| A-20k-Mix | 1,00 | 0,99 | 0,99 | 1,00 | 1,00 | 1,00 | 1,00 | 0,00 | 0,98 | 0,98 | 0,87 |
| A-20k-RBF | 0,68 | 0,71 | 0,69 | 0,55 | 0,73 | 0,72 | 0,75 | 0,04 | 0,65 | 0,65 | 0,54 |
| A-20k-Sine | 1,00 | 0,99 | 0,99 | 0,99 | 1,00 | 1,00 | 1,00 | 0,06 | 0,98 | 0,98 | 0,80 |
| A-20k-Wav | 0,71 | 0,65 | 0,57 | 0,48 | 0,66 | 0,67 | 0,66 | 0,07 | 0,67 | 0,67 | 0,53 |
| A-130k-Agr | 0,62 | 0,90 | 0,85 | 0,58 | 0,89 | 0,87 | 0,91 | 0,01 | 0,61 | 0,61 | 0,56 |
| A-130k-LED | 0,21 | 0,17 | 0,08 | 0,11 | 0,26 | 0,26 | 0,21 | 0,00 | 0,14 | 0,14 | 0,23 |
| A-130k-Mix | 1,00 | 0,99 | 0,99 | 0,95 | 1,00 | 1,00 | 1,00 | 0,00 | 0,98 | 0,98 | 0,79 |
| A-130k-RBF | 0,34 | 0,58 | 0,61 | 0,26 | 0,59 | 0,54 | 0,64 | 0,09 | 0,33 | 0,33 | 0,32 |
| A-130k-Sine | 0,99 | 0,98 | 0,99 | 0,72 | 0,99 | 0,99 | 0,99 | 0,08 | 0,97 | 0,97 | 0,76 |
| A-130k-Wav | 0,66 | 0,72 | 0,63 | 0,29 | 0,69 | 0,68 | 0,71 | 0,04 | 0,65 | 0,65 | 0,46 |
| A-890k-Agr | 0,29 | 0,83 | 0,84 | 0,26 | 0,79 | 0,69 | 0,84 | 0,01 | 0,28 | 0,28 | 0,28 |
| A-890k-LED | 0,10 | 0,12 | 0,05 | 0,05 | 0,17 | 0,16 | 0,14 | 0,00 | 0,07 | 0,07 | 0,08 |
| A-890k-Mix | 1,00 | 0,98 | 0,98 | 0,59 | 1,00 | 1,00 | 1,00 | 0,00 | 0,98 | 0,98 | 0,76 |
| A-890k-RBF | 0,14 | 0,36 | 0,46 | 0,12 | 0,37 | 0,31 | 0,41 | 0,03 | 0,14 | 0,14 | 0,14 |
| A-890k-Sine | 0,92 | 0,92 | 0,94 | 0,40 | 0,95 | 0,95 | 0,96 | 0,04 | 0,88 | 0,88 | 0,65 |
| A-890k-Wav | 0,34 | 0,47 | 0,57 | 0,12 | 0,46 | 0,45 | 0,49 | 0,01 | 0,31 | 0,31 | 0,34 |
| Vitórias | 5 | 5 | 3 | 1 | 7 | 6 | 8 | 0 | 3 | 3 | 1 |
| Derrotas | 3 | 2 | 1 | 8 | 0 | 0 | 0 | 10 | 5 | 5 | 8 |
| Média Postos | 4,25 | 4,94 | 5,64 | 8,89 | 3,00 | 3,31 | 2,75 | 11,00 | 7,03 | 7,03 | 8,17 |

Fonte: O AUTOR (2020).

4.7 CONSIDERAÇÕES FINAIS

Esta seção apresentou os métodos DDE, MSDDE e MDDE, inspirados em métodos comitês de classificadores (FREUND; SCHAPIRE et al., 1996; OZA, 2001; OZA; RUSSELL, 2001). Os critérios de decisões dos métodos de comitês proposto neste trabalho são consideravelmente simples. No DDE e MSDDE, há uma combinação direta dos sinais de detecções para obter os resultados. Apesar da simplicidade dos algoritmos, eles têm-se mostrado eficientes, principalmente para reduzir falsos alarmes. O MDDE, apesar de trabalhar também combinando os sinais de detecções, possui uma característica a mais para calcular a similaridade entre sinais de detecções.

Os métodos DDE, MSDDE e MDDE foram testados com dois algoritmos de aprendizado diferentes, o *Hoeffding Tree* (HT) e *Naive Bayes* (NB). Evidências empíricas mostram que os métodos propostos apresentam resultados satisfatórios em relação as métricas de desempenho utilizadas e membros utilizados nos comitês. Além disso, também se mostrou eficiente em relação aos demais métodos de comitês avaliados.

Por fim, os resultados da métrica MCC foram utilizados para apresentar os falsos alarmes. Houve um ganho de eficiência no desempenho dos métodos propostos em relação aos avaliados, considerando a média entre os resultados obtidos com HT e NB. Em relação aos melhores resultados dos métodos avaliados de única análise e de múltipla análise, o melhor dos métodos de única (FHDDM) obteve aproximadamente 57% do desempenho do MDDE e nenhum dos métodos de múltipla análise foi superior estatisticamente a qualquer um dos métodos avaliados. Finalmente, entre os métodos propostos, o MDDE, seguido do DDE obtiveram os melhores resultados.

5 DETECTORES DE ÚNICA ANÁLISE

Nesta seção, são apresentados dois novos detectores de única análise: *Ultimately Simple Drift Detector* (USDD) e *USDD eXtension* (USDDX). Eles possuem baixas taxas de falsos alarmes, complexidade computacional $\Theta(1)$ e mostraram nos resultados empíricos uma competitiva acurácia da classificação, apesar da simplicidade da estratégia de DMC.

Segundo Barros e Santos (2018), Gonçalves Jr. et al. (2014), um algoritmo de DMC (detector) é um componente de software responsável por analisar os resultados da classificação do fluxo de dados e estimar mudanças de conceito para permitir adaptações no modelo de aprendizado em tempo real. Para ser eficaz, o algoritmo de aprendizado precisa evoluir seu modelo preditivo para que ele siga as mudanças e se adapte rapidamente a elas. Logo, a MC é uma questão fundamental na mineração de fluxo de dados porque os modelos gerados precisam ser adaptados quando ocorre uma MC para permitir que haja garantias teóricas em relação ao aprendizado.

Várias abordagens de DMC já foram propostas na literatura, a exemplo do: DDM, EDDM, ECDD, FHDDM, FHDDMS, FSDD, FPDD, FTDD, HDDMA, HDDMW, RDDM, SEED, STEPD, SeqDrift1, SeqDrift2 e WSTD. Todos esses métodos foram apresentados na seção 2 e seus desempenhos foram mostrados no apêndice B, sendo possível observar a diferença de desempenho entre eles nas métricas observadas. Apesar de haver múltiplas abordagens, há pontos fortes e fracos em mesmo cenário de utilização, quando comparadas seus desempenhos em termos de acurácia da classificação e falso alarmes de detecção, resultando consequências como: efeito limitado nos resultados de acurácia da classificação; elevada demanda e/ou complexidade computacional; baixa sensibilidade à DMC gradual e/ou alta taxa de falsos alarmes.

No apêndice D, as Tabelas 92 e 93 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de *Wilcoxon Signed-Rank*) com as médias de acurácias dos métodos testados nos conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. As Tabelas 96 e 97 apresentam os resultados dos testes das hipóteses estatísticas com as médias de MCC dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Seguidamente, as Tabelas 94 e 95 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de *Wilcoxon Signed-Rank*) com as médias de MTR dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Observe que, para cada uma das três hipóteses H_0 , H_1 e H_2 testadas, na mesma linha, o melhor resultado é mostrado em **negrito**. Essas tabelas apresentam a informação de qual método foi igual (empate), superior (vitória) ou inferior (derrota) estatisticamente a qualquer método avaliado. Na avaliação, as derrotas são subtraídas das vitórias para obter o rank (das diferenças).

5.1 USDD: ULTIMATELY SIMPLE CONCEPT DRIFT DETECTION

Esta seção fornece uma descrição detalhada do USDD, o algoritmo de DMC proposto. Ele se baseia na inferência estatística e no pressuposto de que, dentro de um único conceito, a taxa de erro do classificador tende a diminuir devido à generalização dos casos até atingir um certo nível em que se torna estável (CORTES; VAPNIK, 1995; MITCHELL, 1997). Logo, assume-se como estratégia que qualquer perturbação causadora de degradação a nível significativo na taxa de erro da classificação é considerado uma MC.

Para detectar as mudanças, o USDD realiza uma avaliação do desempenho do classificador. Para isso, ele mantém uma amostra de dados de tamanho n com acertos e erros recentes da classificação. A cada nova amostra, a taxa de erro é recalculada e, quando ela alcança os níveis de *warning* (α_w) e *drift* (α_d), o estado da classificação do detector muda. Consequentemente, até certo ponto, o USDD funciona de forma semelhante ao DDM e demais detectores, porque também monitora as previsões do algoritmo de aprendizado usando amostra de dados.

O USDD não depende da identificação da distribuição de probabilidade dos dados e assume que as previsões feitas pelo classificador são independentes umas das outras. Entende-se também que a distribuição da soma (ou média) de um grande número de variáveis aleatórias Independent and Identically Distributed (traduzindo para o português, Independente e Identicamente Distribuída) (IID) é aproximadamente normal, independentemente da distribuição subjacente dessas variáveis (SEJDIC; FALK, 2018). Espera-se que esta distribuição convirja de qualquer distribuição para uma distribuição normal padrão quando o número de exemplos tende ao infinito.

É importante mencionar que alguns dos métodos de DMC são baseados em janelas (por exemplo, STEP, FTDD, WSTD etc.) e comparam as distribuições de duas janelas de dados, uma com os exemplos mais recentes e outra com informações históricas, possivelmente levando a mais uso de memória. Um dos principais pontos fortes do USDD está na simplicidade de sua estratégia para DMC. Basicamente, ele monitora o desempenho do classificador em tempo real e assume que uma MC ocorre apenas quando a precisão do classificador na amostra de dados cai abaixo de 50%, ou seja, os erros excedem as respostas corretas.

Estratégia semelhante é adotada em vários algoritmos de comitê, por exemplo, AdBoost.M1 (FREUND, 1995; FREUND; SCHAPIRE et al., 1996), *Boosting-like Online Learning Ensemble* (BOLE) (BARROS; SANTOS; GONÇALVES JR., 2016), *Online AdaBoost-based M1* (OABM1) (SANTOS; BARROS, 2019), etc., onde os classificadores fracos com erros médios maiores que 1/2 são descartados, porque são considerados prejudiciais ao comitê.

Um dos parâmetros do USDD é o tamanho da janela deslizante ($w = 80$). O parâmetro r é responsável por determinar o número mínimo de dados que devem ser ignorados

antes que o método passe a detectar possíveis mudanças de conceito. Há também os níveis de erros (porcentagem) para a detecção de warning ($\alpha_w = 0,35$) e drift ($\alpha_d = 0,46$). Esses valores específicos foram escolhidos empiricamente para maximizar a acurácia do classificador. Foram feitos experimentos com parametrizações exploratórias utilizando geradores de dados artificiais e conjuntos de dados reais. Para otimizar as DMC, os valores sugeridos são $w = 80$, $r = 1000$, $\alpha_w = 0,4$ e $\alpha_d = 0,5$. Também é importante observar que conceitos maiores contribuem para induzir melhores generalizações, minimizando os erros, e que tais cenários tornam a escolha dos valores dos parâmetros menos relevante para os resultados de classificação e detecção. Sendo assim, para melhorar a precisão das detecções o limiar de 50% se mostrou eficiente, por outro lado houve uma pequena redução na reação de detecção da MC, fazendo com que o modelo demore um pouco mais para ser construído, implicando em redução na quantidade de treinamento e acurácia da classificação.

Apesar da simplicidade, sua eficiência é sustentada por propriedades teóricas da inferência estatística. Com base no modelo de aprendizado PAC (MITCHELL, 1997), o número de erros do classificador tende a ser estável ou diminuir com mais exemplos dentro de um conceito estacionário, e um aumento significativo no número de erros sugere mudança na distribuição dos dados, ou seja, MC.

O algoritmo 4, apresenta o pseudo-código que foi implementado em linguagem de programação Java. Entre as linhas 18 e 32 estão os critérios para tomadas de decisão em relação a classificação do fluxo de dados. Na linha 19, é testada a condição de estabilidade do fluxo, caso não seja atendida, é avaliada a condição da linha 23, que verifica a ocorrência de *warning*, em caso de ser falsa a condição, o fluxo é classificado como *drift*.

Algoritmo 4: USDD - Ultimately Simple Drift Detector

Input:
stream: data stream of examples, *w*: window size, α_w : warning level, α_d : drift level, *n*: number of examples ignore

```

1 warnLimit  $\leftarrow w \times \alpha_w$  //Quantidade de erros necessários para sinalizar warning
2 driftLimit  $\leftarrow w \times \alpha_d$  //Quantidade de erros necessários para sinalizar drift
3 numIgnoredInst  $\leftarrow n - w$  //Quantidade de respostas ignoradas no início de um novo conceito
4 numInst  $\leftarrow 0$  //Quantidade de instâncias processadas no conceito atual
5 numErr  $\leftarrow 0$  //Quantidade de erros armazenados na janela deslizando
6 pos  $\leftarrow 0$  //Índice utilizado na janela deslizando
7 stPred [*]  $\leftarrow 0$ 

8 foreach instance in stream do
9   pred  $\leftarrow$  prediction (instance) //Resposta de acerto ou erro da classificação
10  numInst  $\leftarrow$  numInst + 1
11  if numInst > numIgnoredInst then
12    numErr  $\leftarrow$  numErr - stPred [pos] + pred //Atualiza quantidade de erros
13    stPred [pos]  $\leftarrow$  pred //A resposta de acerto ou erro da classificação é armazenada na janela
    deslizando
14    pos  $\leftarrow$  pos + 1 //Atualiza a posição do índice da janela deslizando
15    if pos = w then
16      pos  $\leftarrow$  0
17    end
18    if numInst  $\geq$  n then
19      if numErr < warnLimit then
20        raise stable
21      end
22      else
23        if numErr < driftLimit then
24          raise warning
25        end
26        else
27          numInst, numErr, pos  $\leftarrow$  0
28          stPred [*]  $\leftarrow$  0
29          raise drift
30        end
31      end
32    end
33  end
34 end

```

Fonte: O AUTOR (2020).

5.2 USDDX: USDD EXTENSION

Esta seção fornece uma descrição do USDDX, o algoritmo de parametrização dinâmica do USDD. A proposta do USDDX é de otimizar os resultados do USDD e para isso ele ajusta seus parâmetros de configuração dinamicamente, levando em consideração as condições do fluxo de dados. O ajuste ocorre no parâmetro de tamanho da janela deslizando usada para inferir a amplitude da onda e supor estabilidade do conceito aprendido.

Na parametrização dinâmica, são adotados dois níveis para ajustes. No primeiro, assume que o conceito aprendido possui acurácia média maior que 90%; no segundo, maior que 85%; e, para os demais valores, assume-se que os parâmetros por padrão são satisfatórios. A suposição levada em consideração para realizar tais ajustes é com base nas explicações teóricas do teorema central do limite, pois, em uma situação ideal (distribuição normal padrão), a média amostral subtraída da média populacional será zero e a variância 1. Assim sendo, por inferência estatística em um fluxo de dados com acurácia média maior que 85%, será possível assumir uma quantidade menor de observações para supor a média amostral e tomar uma decisão acerca das mais rapidamente.

Quando a acurácia média for maior que 85% ou 90%, redimensiona-se o tamanho da janela para 1/2 ou 1/4 do valor definido, respectivamente. A acurácia média é calculada com base na soma de erros fornecidos pela janela deslizando de tamanho *w*. Após três vezes

o período assumido pelo USDD para que o aprendizado do conceito seja “estabilizado”, são somados todos os erros médios obtidos da janela w durante o mesmo tamanho de período aguardado para estabilizar o conceito.

Uma vez que a acurácia média enquadra-se em um dos dois níveis mencionados, é preciso ajustar o tamanho da janela e critérios de decisão. O teste de DMC não muda em relação ao USDD a mudança apenas altera os parâmetros do método. Por fim, a ideia geral do USDDX é otimizar os parâmetros de acordo com o contexto, fazendo uma espécie de sobreajuste para o conceito atual.

O algoritmo 5 apresenta o pseudo-código que foi implementado em linguagem de programação Java. Entre as linhas 48 e 61 estão os critérios para tomadas de decisão em relação à classificação do fluxo de dados. Na linha 48, é testada a condição de estabilidade do fluxo, caso não seja atendida é avaliada a condição da linha 52, que verifica a ocorrência de *warning*, em caso de ser falsa a condição, o fluxo é classificado como *drift*.

O algoritmo 5 tem como entrada as variáveis *stream*, dw , α_w , α_d e n . O fluxo de dados é referenciado pela variáveis *stream*, dw indica o tamanho padrão da janela, α_w e α_d indicam valores que vão de 0 a 1 para limites de ocorrências de *warning* e *drift*, respectivamente. Por fim, n define a quantidade de predições feitos classificador que devem ser ignoradas antes do início das detecções.

As variáveis w , *warniLimit*, *driftLimit* e *numIgnoredInst* são inicializadas com base nos parâmetros do algoritmos. A variável w inicialmente possui o valor de dw , *warningLimit* e *driftLimit* são definidas com base no produto de w e valores de α_w e α_d , respectivamente. Finalmente, *numIgnoredInst* é calculada pela subtração de w em n , sendo esta a quantidade de classificações que serão ignoradas antes de preencher a janela deslizante com acertos e erros do modelo.

A janela deslizante é composta pelas variáveis *stPred*, *pos* e w . A janela implementa o conceito de lista circular, no qual as respostas de classificações antigas são descartadas para dá lugar a novas. O índice de posição da lista é representado por *pos*, o tamanho da lista é dado por w , e, finalmente, as respostas são guardados no vetor de valores *stPred*.

A quantidade de elementos recebidos para análise é indicado por *numInst*. Em *numErr* é registrado o valor atual da quantidade de erros do classificado dentro da janela deslizante implementado por *stPred*. A variável *pred* possui a resposta de acerto ou erro da classificação.

Na linha 6, *dSize* recebe o valor falso na inicialização. A variável *dSize* é utilizada como controladora do mecanismo de estimação do tamanho da janela em tempo de execução. Para cada fluxo de dados, é estimado uma única vez o tamanho da janela deslizante. Essa estimativa leva em consideração a quantidade de erros na classificação atual, calculadas por *muAdjust* na linha 23. A variável *rw* é inicialmente igual a w e em caso de ajuste no tamanho da janela, *rw* é modificada para o novo tamanho.

Algoritmo 5: USDDX - Ultimately Simple Drift Detector eXtension**Input:**

stream: data stream of examples, *dw*: default window size, α_w : warning level, α_d : drift level, *n*: number of examples
ignore

```

1  w ← dw
2  warnLimit ← w ×  $\alpha_w$ 
3  driftLimit ← w ×  $\alpha_d$ 
4  numIgnoredInst ← n − w
5  numInst, numAjust, sumAjust, numErr, pos, stPred [*] ← 0
6  dSize ← false
7  foreach instance in stream do
8    pred ← prediction(instance) //Resposta de acerto ou erro da classificação
9    numInst ← numInst + 1
10   if numInst > numIgnoredInst then
11     numErr ← numErr − stPred[pos] + pred //Atualiza quantidade de erros
12     stPred[pos] ← pred //Armazena resposta da classificação
13     pos ← pos + 1 //Atualiza a posição do índice
14     if pos = w then
15       pos ← 0
16     if numInst ≥ n then
17       if dSize = false then
18         if numInst > n * 3 then
19           numAjust ← numAjust + 1
20           if numAjust < n then
21             sumAjust ← sumAjust + numErr
22           else
23             muAjust ← 1 − ((sumAjust / n) / 100)
24             rw ← w
25             if muAjust > 0.9 then
26               rw ← w * 0.25
27             else
28               if muAjust > 0.85 then
29                 rw ← w * 0.5
30             end
31             if not rw = w then
32               w ← rw
33               warnLimit ← w ×  $\alpha_w$ 
34               driftLimit ← w ×  $\alpha_d$ 
35               stPred [*] ← 0
36               dSize ← true
37             end
38             if numErr < warnLimit then
39               raise stable
40             else
41               if numErr < driftLimit then
42                 raise warning
43               else
44                 w ← dw
45                 warnLimit ← w ×  $\alpha_w$ 
46                 driftLimit ← w ×  $\alpha_d$ 
47                 numInst, numErr, pos ← 0
48                 stPred [*] ← 0
49                 dSize ← false
50                 raise drift
51             end
52         end
53 end

```

Fonte: O AUTOR (2020).

5.3 RESULTADOS E DISCUSSÃO

As Tabelas 10 e 11 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Os cenários avaliados nessas tabelas incluem três tamanhos de base de dados com simulações de mudanças abruptas e graduais, bem como em conjuntos de dados do mundo real. É importante destacar que, para cada conjunto de dados, na mesma linha, o melhor resultado é mostrado em **negrito**. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos que obtiveram os melhores resultados com HT e NB foram o USDD e USDDX, respectivamente.

A Figura 10 (a) apresenta o resumo dos resultados com HT e NB apresentados no apêndice E nas Tabelas 92 e 93. Em termos das hipóteses testadas, o USDD e USDDX

com HT forneceram os maiores números de vitórias, 7 e 6, respectivamente, e nenhuma derrota. Já com NB o melhor foi o USDDX com 8 vitórias, seguido do USDD com 6 vitórias e sem derrotas. O método USDD foi igual estatisticamente aos métodos FTDD, FSDD e FPDD com HT. No caso do NB, os detectores equivalentes em termos estatísticos ao USDDX foram o USDD e o RDDM. O RDDM, apesar de ter obtido um resultado positivo com o NB (3), não conseguiu manter o mesmo ou melhorar o desempenho com o HT, alcançando uma diferença de -2 . Os métodos com maior número de derrotas usando o HT foram HDDMW e WSTD, ambos com 4 derrotas e com o NB foram FTDD (8), WSTD (6) e FTDD (4).

É fácil notar que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos métodos USDD e USDDX, todos com avaliação acima de zero, como pode ser visto na Figura 10 (b). Apesar disso, vale a pena destacar que o desempenho do USDD e USDDX foi sólido na maioria das situações e não houve inferioridade estatística em relação a outros com o HT e NB. A Figura 10 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB, logo, é possível observar que os métodos USDD e USDDX obtiveram as maiores médias (6, 5) no desempenho acumulado, seguidos pelo FSDD com 1, RDDM (0, 5), FHDDM (0), HDDMA ($-0, 5$), FHDDMS e FTDD com $-1, 5$, HDDMW (-2), FSDD (-4) e WSTD com -5 . Vale ressaltar que apenas os métodos FHDDM, FSDD e RDDM obtiveram rank com valor positivo (acima de zero), entretanto, ainda ficaram com valores numéricos muito abaixo dos obtidos pelos métodos USDD e USDDX.

Tabela 10 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _b | RDDM | WSTD | USDD | USDDX |
|--------------|-------|--------|-------|-------|-------|-------------------|-------------------|-------|-------|-------|-------|
| A-20k-Agr | 86,92 | 86,95 | 86,20 | 86,44 | 86,20 | 86,67 | 86,92 | 86,80 | 86,99 | 86,73 | 87,04 |
| A-20k-LED | 69,98 | 70,06 | 68,55 | 68,48 | 68,54 | 70,02 | 69,91 | 70,12 | 68,24 | 70,45 | 70,45 |
| A-20k-Mix | 90,75 | 90,75 | 90,82 | 90,82 | 90,82 | 90,49 | 90,82 | 90,76 | 90,81 | 90,46 | 90,77 |
| A-20k-RBF | 74,83 | 74,78 | 74,72 | 74,94 | 74,81 | 74,79 | 74,73 | 74,77 | 74,96 | 74,90 | 74,92 |
| A-20k-Sine | 89,73 | 89,73 | 89,76 | 89,76 | 89,76 | 89,67 | 89,80 | 89,55 | 89,77 | 89,45 | 89,73 |
| A-20k-Wav | 79,32 | 79,36 | 79,34 | 79,38 | 79,37 | 79,29 | 79,36 | 79,31 | 79,39 | 79,37 | 79,29 |
| A-130k-Agr | 91,50 | 91,39 | 92,02 | 92,03 | 92,02 | 91,74 | 90,85 | 91,72 | 91,90 | 91,55 | 91,47 |
| A-130k-LED | 71,45 | 71,42 | 71,02 | 70,74 | 71,06 | 71,47 | 71,29 | 71,52 | 70,26 | 71,66 | 71,66 |
| A-130k-Mix | 92,60 | 92,60 | 92,58 | 92,58 | 92,58 | 92,57 | 92,62 | 92,29 | 92,59 | 92,56 | 92,48 |
| A-130k-RBF | 77,99 | 77,47 | 79,61 | 79,59 | 79,61 | 79,27 | 76,67 | 79,35 | 79,03 | 78,84 | 78,81 |
| A-130k-Sine | 93,34 | 93,34 | 93,34 | 93,34 | 93,34 | 93,28 | 93,32 | 92,99 | 93,32 | 93,28 | 93,28 |
| A-130k-Wav | 79,62 | 79,76 | 79,57 | 79,60 | 79,59 | 79,57 | 79,87 | 79,72 | 79,62 | 79,57 | 79,69 |
| A-890k-Agr | 95,16 | 95,12 | 95,26 | 95,26 | 95,26 | 95,14 | 94,91 | 93,92 | 95,19 | 95,18 | 95,15 |
| A-890k-LED | 71,71 | 71,68 | 71,72 | 71,25 | 71,70 | 71,71 | 71,58 | 71,74 | 70,63 | 71,90 | 71,90 |
| A-890k-Mix | 94,97 | 94,97 | 94,95 | 94,95 | 94,95 | 94,95 | 94,97 | 93,89 | 94,96 | 94,97 | 94,90 |
| A-890k-RBF | 83,55 | 81,80 | 85,18 | 85,09 | 85,16 | 84,37 | 79,46 | 83,28 | 83,83 | 84,89 | 84,88 |
| A-890k-Sine | 96,97 | 96,97 | 96,96 | 96,96 | 96,96 | 96,91 | 96,96 | 95,67 | 96,94 | 96,97 | 96,96 |
| A-890k-Wav | 81,22 | 80,07 | 82,19 | 81,97 | 82,17 | 81,76 | 79,97 | 80,42 | 81,13 | 82,23 | 82,03 |
| G-20k-Agr | 85,17 | 85,47 | 83,89 | 84,32 | 83,92 | 85,30 | 85,28 | 85,48 | 83,44 | 85,16 | 85,40 |
| G-20k-LED | 69,65 | 69,66 | 68,40 | 68,17 | 68,49 | 69,54 | 69,51 | 69,71 | 68,04 | 70,00 | 70,00 |
| G-20k-Mix | 87,77 | 87,42 | 87,18 | 87,25 | 87,22 | 87,48 | 87,54 | 87,59 | 87,45 | 87,73 | 87,50 |
| G-20k-RBF | 73,91 | 73,72 | 73,70 | 74,04 | 73,78 | 74,13 | 73,79 | 73,98 | 73,14 | 74,21 | 74,23 |
| G-20k-Sine | 86,99 | 86,75 | 86,44 | 86,53 | 86,47 | 86,74 | 86,84 | 86,89 | 86,73 | 86,97 | 86,79 |
| G-20k-Wav | 78,79 | 78,89 | 78,17 | 78,60 | 78,40 | 78,66 | 78,89 | 79,03 | 78,53 | 78,65 | 78,76 |
| G-130k-Agr | 91,19 | 91,13 | 91,72 | 91,74 | 91,72 | 91,51 | 90,58 | 91,57 | 90,99 | 91,27 | 91,18 |
| G-130k-LED | 71,39 | 71,35 | 70,95 | 70,76 | 70,98 | 71,36 | 71,23 | 71,45 | 70,16 | 71,58 | 71,58 |
| G-130k-Mix | 92,14 | 92,08 | 92,03 | 92,03 | 92,03 | 92,06 | 92,11 | 92,07 | 92,05 | 92,14 | 91,95 |
| G-130k-RBF | 77,68 | 77,24 | 79,32 | 79,38 | 79,41 | 79,15 | 76,49 | 79,37 | 78,49 | 78,51 | 78,44 |
| G-130k-Sine | 92,84 | 92,82 | 92,77 | 92,77 | 92,77 | 92,77 | 92,84 | 92,75 | 92,83 | 92,86 | 92,80 |
| G-130k-Wav | 79,68 | 79,65 | 79,75 | 79,63 | 79,84 | 79,59 | 79,70 | 79,75 | 79,43 | 79,79 | 79,75 |
| G-890k-Agr | 95,10 | 95,08 | 95,23 | 95,23 | 95,23 | 95,10 | 94,86 | 93,89 | 94,99 | 95,12 | 95,11 |
| G-890k-LED | 71,70 | 71,68 | 71,70 | 71,23 | 71,69 | 71,72 | 71,57 | 71,74 | 70,60 | 71,89 | 71,89 |
| G-890k-Mix | 94,90 | 94,88 | 94,86 | 94,85 | 94,86 | 94,88 | 94,88 | 93,86 | 94,87 | 94,90 | 94,82 |
| G-890k-RBF | 83,23 | 81,64 | 85,11 | 85,07 | 85,09 | 84,30 | 79,51 | 83,22 | 83,71 | 84,77 | 84,71 |
| G-890k-Sine | 96,89 | 96,87 | 96,85 | 96,84 | 96,84 | 96,85 | 96,88 | 95,73 | 96,86 | 96,88 | 96,87 |
| G-890k-Wav | 81,04 | 79,97 | 82,02 | 81,64 | 82,00 | 81,66 | 79,89 | 80,36 | 81,01 | 82,04 | 81,99 |
| Airlines | 65,73 | 65,81 | 65,80 | 65,71 | 65,77 | 66,00 | 65,43 | 66,54 | 65,93 | 65,18 | 65,17 |
| Connect4 | 75,21 | 75,14 | 74,41 | 74,88 | 74,57 | 75,14 | 75,21 | 75,18 | 75,18 | 74,97 | 74,92 |
| Outdoor | 60,09 | 60,25 | 59,66 | 58,46 | 59,21 | 59,96 | 60,85 | 58,63 | 60,12 | 60,88 | 60,88 |
| Rialto | 41,75 | 47,29 | 36,62 | 37,10 | 33,45 | 44,24 | 45,28 | 44,52 | 38,77 | 55,44 | 55,44 |
| SpamData | 92,57 | 92,92 | 92,18 | 92,18 | 92,18 | 92,76 | 92,46 | 92,30 | 92,43 | 91,53 | 91,66 |
| WineRed | 52,07 | 52,07 | 53,15 | 52,71 | 53,17 | 52,82 | 52,50 | 52,76 | 51,34 | 53,60 | 53,60 |
| WineWhite | 46,09 | 46,10 | 45,21 | 44,44 | 44,44 | 45,15 | 46,71 | 44,32 | 45,02 | 46,98 | 46,98 |
| Vitórias | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 7 | 6 |
| Derrotas | 2 | 2 | 0 | 0 | 0 | 2 | 4 | 2 | 4 | 0 | 1 |
| Média Postos | 5,17 | 6,10 | 6,42 | 6,45 | 6,15 | 6,28 | 6,28 | 6,44 | 6,98 | 4,37 | 5,35 |

Fonte: O AUTOR (2020).

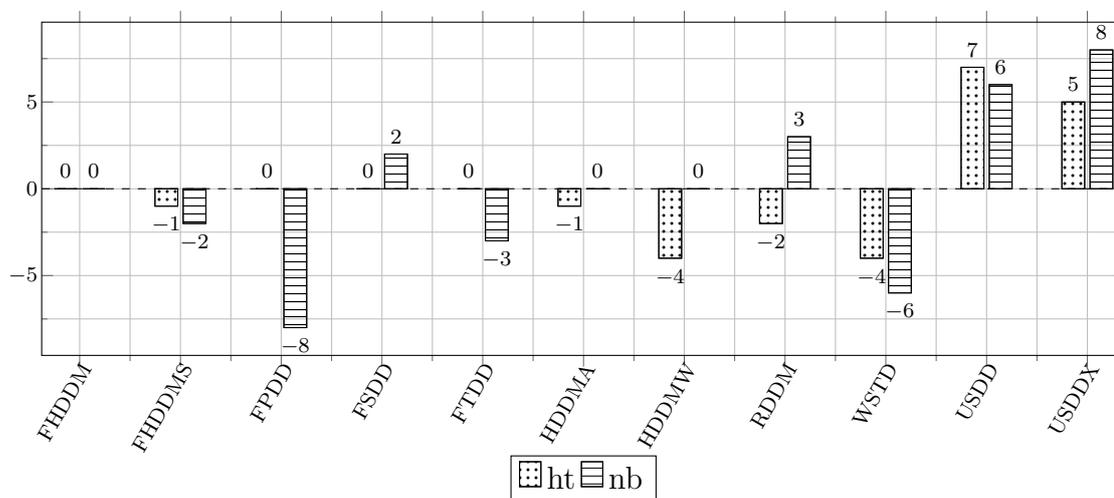
Tabela 11 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | WSTD | USDD | USDDX |
|--------------|--------------|--------------|--------------|--------------|--------------|-------------------|-------------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,95 | 86,99 | 85,71 | 86,83 | 85,71 | 86,86 | 86,94 | 86,89 | 86,99 | 86,77 | 87,07 |
| A-20k-LED | 71,04 | 71,08 | 66,15 | 67,97 | 66,10 | 71,40 | 70,56 | 71,83 | 68,06 | 71,53 | 71,53 |
| A-20k-Mix | 91,11 | 91,11 | 91,19 | 91,19 | 91,19 | 91,10 | 91,19 | 91,03 | 91,19 | 90,81 | 91,23 |
| A-20k-RBF | 74,88 | 74,91 | 74,76 | 74,89 | 74,83 | 74,77 | 74,90 | 74,72 | 74,84 | 74,89 | 74,89 |
| A-20k-Sine | 89,15 | 89,15 | 89,17 | 89,18 | 89,18 | 89,04 | 89,20 | 88,99 | 89,19 | 88,88 | 89,18 |
| A-20k-Wav | 79,70 | 79,78 | 79,43 | 79,69 | 79,55 | 79,46 | 79,81 | 79,63 | 79,66 | 79,56 | 79,65 |
| A-130k-Agr | 88,71 | 88,70 | 88,90 | 88,91 | 88,90 | 88,75 | 88,64 | 88,77 | 88,89 | 88,82 | 88,83 |
| A-130k-LED | 72,63 | 72,53 | 70,97 | 72,62 | 70,99 | 73,44 | 72,14 | 73,48 | 72,03 | 73,45 | 73,45 |
| A-130k-Mix | 91,93 | 91,93 | 91,94 | 91,94 | 91,94 | 91,84 | 91,94 | 91,84 | 91,94 | 91,88 | 91,94 |
| A-130k-RBF | 75,26 | 75,24 | 75,28 | 75,31 | 75,28 | 75,25 | 75,18 | 75,22 | 75,21 | 75,30 | 75,30 |
| A-130k-Sine | 89,48 | 89,48 | 89,49 | 89,49 | 89,49 | 89,37 | 89,49 | 89,37 | 89,49 | 89,44 | 89,50 |
| A-130k-Wav | 80,33 | 80,32 | 80,21 | 80,33 | 80,30 | 80,22 | 80,30 | 80,21 | 80,29 | 80,32 | 80,31 |
| A-890k-Agr | 89,10 | 89,07 | 89,37 | 89,36 | 89,37 | 89,30 | 89,01 | 89,25 | 89,29 | 89,26 | 89,25 |
| A-890k-LED | 73,02 | 72,89 | 73,71 | 73,82 | 73,70 | 73,87 | 72,49 | 73,84 | 73,36 | 73,83 | 73,83 |
| A-890k-Mix | 92,07 | 92,07 | 92,08 | 92,08 | 92,08 | 92,03 | 92,08 | 92,02 | 92,08 | 92,07 | 92,06 |
| A-890k-RBF | 75,36 | 75,31 | 75,44 | 75,46 | 75,46 | 75,43 | 75,23 | 75,40 | 75,40 | 75,43 | 75,43 |
| A-890k-Sine | 89,50 | 89,50 | 89,50 | 89,50 | 89,50 | 89,45 | 89,50 | 89,54 | 89,50 | 89,49 | 89,51 |
| A-890k-Wav | 80,38 | 80,36 | 80,40 | 80,40 | 80,40 | 80,35 | 80,33 | 80,37 | 80,39 | 80,41 | 80,40 |
| G-20k-Agr | 84,07 | 84,90 | 80,47 | 82,95 | 80,48 | 84,79 | 84,70 | 85,07 | 79,54 | 84,20 | 84,85 |
| G-20k-LED | 70,40 | 70,54 | 64,87 | 67,11 | 64,83 | 71,07 | 70,09 | 71,63 | 65,23 | 70,90 | 70,90 |
| G-20k-Mix | 88,18 | 87,85 | 87,64 | 87,67 | 87,65 | 87,81 | 87,99 | 88,00 | 87,86 | 88,11 | 87,82 |
| G-20k-RBF | 73,79 | 73,82 | 73,01 | 73,75 | 73,16 | 73,91 | 73,90 | 73,90 | 69,78 | 74,01 | 73,97 |
| G-20k-Sine | 87,11 | 86,90 | 86,55 | 86,55 | 86,56 | 86,96 | 87,03 | 87,00 | 86,75 | 87,15 | 86,91 |
| G-20k-Wav | 78,95 | 79,09 | 78,88 | 78,93 | 78,92 | 78,97 | 79,10 | 79,16 | 78,47 | 78,94 | 79,01 |
| G-130k-Agr | 88,31 | 88,32 | 88,35 | 88,39 | 88,36 | 88,45 | 88,25 | 88,50 | 86,62 | 88,40 | 88,43 |
| G-130k-LED | 72,54 | 72,44 | 70,36 | 72,11 | 70,35 | 73,41 | 72,06 | 73,47 | 71,17 | 73,35 | 73,35 |
| G-130k-Mix | 91,47 | 91,40 | 91,39 | 91,40 | 91,39 | 91,42 | 91,43 | 91,45 | 91,44 | 91,46 | 91,41 |
| G-130k-RBF | 75,06 | 75,03 | 75,07 | 75,07 | 75,06 | 75,11 | 74,99 | 75,12 | 73,78 | 75,10 | 75,10 |
| G-130k-Sine | 89,31 | 89,31 | 89,17 | 89,19 | 89,18 | 89,28 | 89,32 | 89,31 | 89,25 | 89,33 | 89,30 |
| G-130k-Wav | 80,11 | 80,14 | 79,81 | 80,08 | 80,02 | 80,09 | 80,11 | 80,08 | 79,64 | 80,07 | 80,06 |
| G-890k-Agr | 89,02 | 88,96 | 89,27 | 89,26 | 89,27 | 89,23 | 88,91 | 89,18 | 88,95 | 89,18 | 89,16 |
| G-890k-LED | 73,01 | 72,88 | 73,63 | 73,77 | 73,63 | 73,87 | 72,48 | 73,84 | 73,19 | 73,82 | 73,82 |
| G-890k-Mix | 92,00 | 91,99 | 91,99 | 92,00 | 91,99 | 91,99 | 92,00 | 91,95 | 92,00 | 92,00 | 91,99 |
| G-890k-RBF | 75,34 | 75,29 | 75,41 | 75,42 | 75,42 | 75,42 | 75,20 | 75,38 | 75,18 | 75,40 | 75,40 |
| G-890k-Sine | 89,50 | 89,50 | 89,48 | 89,47 | 89,48 | 89,45 | 89,50 | 89,53 | 89,47 | 89,50 | 89,50 |
| G-890k-Wav | 80,35 | 80,33 | 80,35 | 80,35 | 80,35 | 80,34 | 80,31 | 80,34 | 80,28 | 80,35 | 80,36 |
| Airlines | 65,82 | 65,83 | 66,72 | 66,91 | 66,76 | 67,23 | 65,34 | 67,50 | 66,84 | 65,44 | 65,43 |
| Connect4 | 75,23 | 75,22 | 74,20 | 74,65 | 74,23 | 74,95 | 75,23 | 75,23 | 75,12 | 74,82 | 74,85 |
| Outdoor | 60,59 | 60,63 | 59,77 | 60,23 | 59,80 | 60,11 | 61,97 | 59,85 | 61,03 | 61,19 | 61,19 |
| Rialto | 41,24 | 47,47 | 26,55 | 28,12 | 27,52 | 43,30 | 45,69 | 44,88 | 27,26 | 55,44 | 55,44 |
| SpamData | 91,51 | 91,73 | 91,35 | 91,35 | 91,35 | 90,96 | 91,68 | 91,39 | 91,80 | 91,17 | 91,04 |
| WineRed | 47,46 | 47,46 | 47,70 | 47,85 | 47,70 | 47,63 | 47,54 | 46,92 | 47,85 | 48,22 | 48,22 |
| WineWhite | 46,53 | 46,90 | 44,96 | 45,48 | 44,88 | 43,16 | 47,08 | 43,41 | 45,40 | 47,77 | 47,77 |
| Vitórias | 2 | 0 | 0 | 3 | 1 | 1 | 2 | 3 | 0 | 6 | 8 |
| Derrotas | 2 | 2 | 8 | 1 | 4 | 1 | 2 | 0 | 6 | 0 | 0 |
| Média Postos | 5,99 | 6,22 | 7,48 | 5,42 | 6,76 | 6,40 | 6,00 | 5,47 | 7,02 | 4,79 | 4,47 |

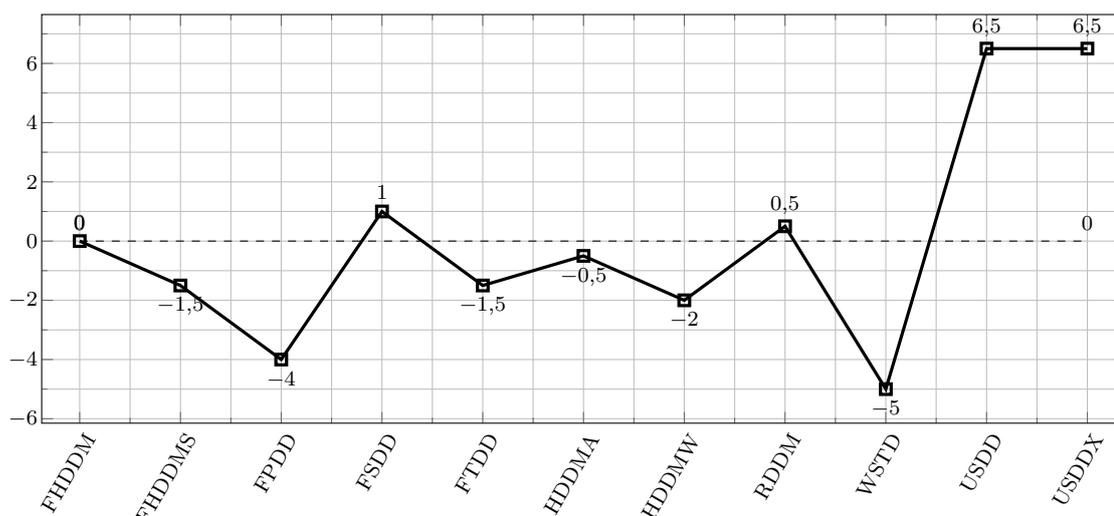
Fonte: O AUTOR (2020).

Figura 10 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



Fonte: O AUTOR (2020).

As Tabelas 12 e 13 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o método USDDX obteve os melhores resultados com HT (5, 11) e NB (4, 5).

A Figura 11 (a) apresenta o resumo dos resultados com HT e NB apresentados no apêndice E nas Tabelas 94 e 95. Em termos das hipóteses testadas, o FSDD, FTDD e WSTD com HT forneceram os melhores desempenhos, ambos com 2 vitórias e nenhuma derrota, já com NB os métodos FSDD e USDDX obtiveram 5 vitórias enquanto o FTDD obteve 4 e também nenhuma derrota.

Entretanto, os métodos FHDDM, HDDMA, HDDMW, WSTD, USDD e USDDX

foram iguais estatisticamente aos métodos FSDD, FTDD e WSTD com HT e com NB HDDMA, HDDMW, WSTD e USDD foram iguais estatisticamente aos métodos FSDD e USDDX. O método com maior número de derrotas usando HT foi o RDDM com 6 derrotas e repete-se com NB com 8 derrotas, logo, tratando-se da ineficiência, ele foi, em termos absolutos e em termos estatísticos, o pior método avaliado. Vale ressaltar que o RDDM não obteve nenhuma superioridade estatística em relação aos demais métodos avaliados.

É fácil notar que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos métodos FSDD, USDDX, FTDD, HDDMW, HDDMA, WSTD e USDD, todos com rank de avaliação acima de zero, mostrados na Figura 11 (b). Apesar disso, vale a pena destacar que o desempenho de FSDD, FTDD e HDDMW foi bom na maioria das situações com HT (2) e com NB foi juntamente líder com USDDX, ambos com 5 vitórias. A Figura 11 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB, logo, é possível observar que o FSDD obteve a maior média (3,5) no desempenho acumulado, seguido pelo USDDX com 3, FTDD (2), HDDMW (1,5), HDDMA, WSTD e USDD com 1 e os demais ficaram com média negativa em número de vitórias.

Tabela 12 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _g | HDDM _w | RDDM | WSTD | USDD | USDDX |
|--------------|--------------|--------------|-----------------|-----------------|-----------------|-------------------|-------------------|---------------|----------------|----------------|----------------|
| A-20k-Agr | 104,44 | 112,43 | 49,29 | 64,39 | 49,07 | 63,50 | 129,13 | 57,86 | 128,64 | 56,10 | 125,56 |
| A-20k-LED | 7,63 | 12,16 | 1,79 | 1,32 | 1,66 | 2,74 | 18,22 | 2,32 | 1,59 | 20,79 | 20,79 |
| A-20k-Mix | 310,41 | 310,41 | 340,77 | 340,77 | 340,77 | 257,89 | 439,48 | 267,77 | 360,21 | 120,34 | 388,27 |
| A-20k-RBF | 32,20 | 41,43 | 2,27 | 7,39 | 4,95 | 14,85 | 49,64 | 18,65 | 11,56 | 14,90 | 15,91 |
| A-20k-Sine | 283,61 | 283,61 | 285,17 | 285,17 | 285,17 | 252,81 | 361,46 | 129,05 | 314,69 | 105,51 | 285,41 |
| A-20k-Wav | 1,47 | 3,65 | 0,78 | 1,08 | 0,84 | 1,20 | 7,19 | 2,16 | 1,15 | 0,90 | 2,16 |
| A-130k-Agr | 732,18 | 805,05 | 961,42 | 1069,78 | 1011,47 | 702,40 | 818,32 | 169,33 | 1177,06 | 327,70 | 659,29 |
| A-130k-LED | 23,39 | 27,43 | 1,70 | 1,45 | 2,12 | 16,49 | 29,12 | 14,57 | 2,86 | 26,16 | 26,16 |
| A-130k-Mix | 1993,70 | 1993,70 | 1970,80 | 1970,80 | 1970,80 | 2497,95 | 2751,53 | 354,40 | 2288,86 | 760,06 | 2770,86 |
| A-130k-RBF | 207,39 | 231,56 | 131,72 | 286,43 | 243,09 | 218,18 | 184,27 | 66,06 | 309,66 | 141,90 | 145,55 |
| A-130k-Sine | 1772,46 | 1772,46 | 2011,33 | 2011,33 | 2011,33 | 2568,86 | 2191,48 | 321,19 | 2142,92 | 661,32 | 1684,97 |
| A-130k-Wav | 63,78 | 75,75 | 1,17 | 23,80 | 8,42 | 21,85 | 34,37 | 21,70 | 18,04 | 12,39 | 26,27 |
| A-890k-Agr | 5475,15 | 5637,86 | 9212,94 | 9069,74 | 9114,79 | 2518,23 | 6194,77 | 267,74 | 8827,81 | 2340,88 | 4405,24 |
| A-890k-LED | 25,54 | 27,96 | 2,72 | 2,34 | 1,63 | 15,16 | 33,81 | 15,53 | 5,03 | 23,36 | 23,35 |
| A-890k-Mix | 13538,84 | 13538,84 | 16451,68 | 16451,68 | 16451,68 | 16981,12 | 17859,12 | 1007,84 | 16339,46 | 5104,01 | 15004,68 |
| A-890k-RBF | 1310,04 | 1129,41 | 2306,12 | 2075,84 | 2412,69 | 1099,40 | 493,44 | 197,84 | 1560,00 | 883,48 | 911,61 |
| A-890k-Sine | 11830,04 | 11830,04 | 16724,42 | 16724,42 | 16724,42 | 13992,72 | 14334,66 | 1250,35 | 15476,77 | 4360,71 | 10606,73 |
| A-890k-Wav | 218,15 | 62,42 | 108,74 | 449,29 | 298,90 | 556,74 | 64,63 | 24,23 | 222,57 | 305,95 | 128,17 |
| G-20k-Agr | 15,47 | 4,59 | 1,59 | 4,08 | 1,64 | 7,43 | 15,11 | 4,24 | 4,55 | 15,39 | 10,59 |
| G-20k-LED | 2,29 | 2,59 | 1,47 | 1,27 | 1,53 | 2,05 | 3,06 | 2,02 | 1,30 | 4,72 | 4,72 |
| G-20k-Mix | 0,95 | 0,91 | 5,01 | 5,49 | 4,98 | 2,68 | 0,92 | 1,32 | 5,98 | 1,07 | 4,75 |
| G-20k-RBF | 6,86 | 6,77 | 1,03 | 1,55 | 1,20 | 7,24 | 3,83 | 9,59 | 0,77 | 5,12 | 4,51 |
| G-20k-Sine | 1,32 | 0,89 | 2,17 | 2,17 | 2,08 | 1,20 | 0,92 | 2,67 | 4,41 | 1,57 | 1,46 |
| G-20k-Wav | 0,64 | 1,14 | 0,18 | 0,39 | 0,28 | 0,47 | 1,33 | 1,67 | 0,41 | 0,34 | 0,44 |
| G-130k-Agr | 129,63 | 43,49 | 70,01 | 107,01 | 96,21 | 99,82 | 69,43 | 153,23 | 90,16 | 119,78 | 116,00 |
| G-130k-LED | 8,38 | 10,88 | 1,55 | 1,73 | 1,71 | 4,93 | 7,35 | 7,92 | 1,59 | 8,40 | 8,45 |
| G-130k-Mix | 0,87 | 0,99 | 28,77 | 32,21 | 29,52 | 45,98 | 5,02 | 95,13 | 36,70 | 156,50 | 55,87 |
| G-130k-RBF | 42,27 | 15,66 | 30,63 | 62,54 | 48,59 | 86,37 | 18,93 | 60,66 | 2,80 | 48,41 | 40,86 |
| G-130k-Sine | 1,50 | 0,79 | 8,27 | 20,95 | 8,42 | 0,87 | 0,85 | 212,37 | 13,15 | 113,31 | 1,44 |
| G-130k-Wav | 3,40 | 2,23 | 0,27 | 6,20 | 1,04 | 9,29 | 2,63 | 9,19 | 0,84 | 2,31 | 2,15 |
| G-890k-Agr | 1091,11 | 1,58 | 661,60 | 661,56 | 661,55 | 538,99 | 1574,17 | 142,82 | 1198,95 | 844,84 | 1112,33 |
| G-890k-LED | 6,91 | 9,01 | 2,12 | 1,22 | 1,65 | 5,38 | 5,38 | 11,40 | 2,34 | 8,56 | 8,12 |
| G-890k-Mix | 173,78 | 0,80 | 303,00 | 294,47 | 308,53 | 0,95 | 57,85 | 116,91 | 381,50 | 1507,10 | 194,92 |
| G-890k-RBF | 241,91 | 16,95 | 588,63 | 526,61 | 573,54 | 289,71 | 59,49 | 136,56 | 87,67 | 330,90 | 281,47 |
| G-890k-Sine | 154,03 | 0,78 | 14,31 | 14,31 | 14,32 | 0,85 | 92,76 | 119,52 | 216,39 | 1975,23 | 218,83 |
| G-890k-Wav | 17,08 | 6,10 | 27,96 | 44,76 | 47,32 | 145,44 | 5,19 | 11,67 | 1,70 | 58,50 | 35,15 |
| Vitórias | 1 | 0 | 0 | 2 | 2 | 1 | 2 | 0 | 1 | 1 | 1 |
| Derrotas | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 |
| Média Postos | 5,58 | 6,58 | 7,19 | 5,78 | 6,44 | 5,81 | 5,14 | 6,69 | 5,78 | 5,89 | 5,11 |

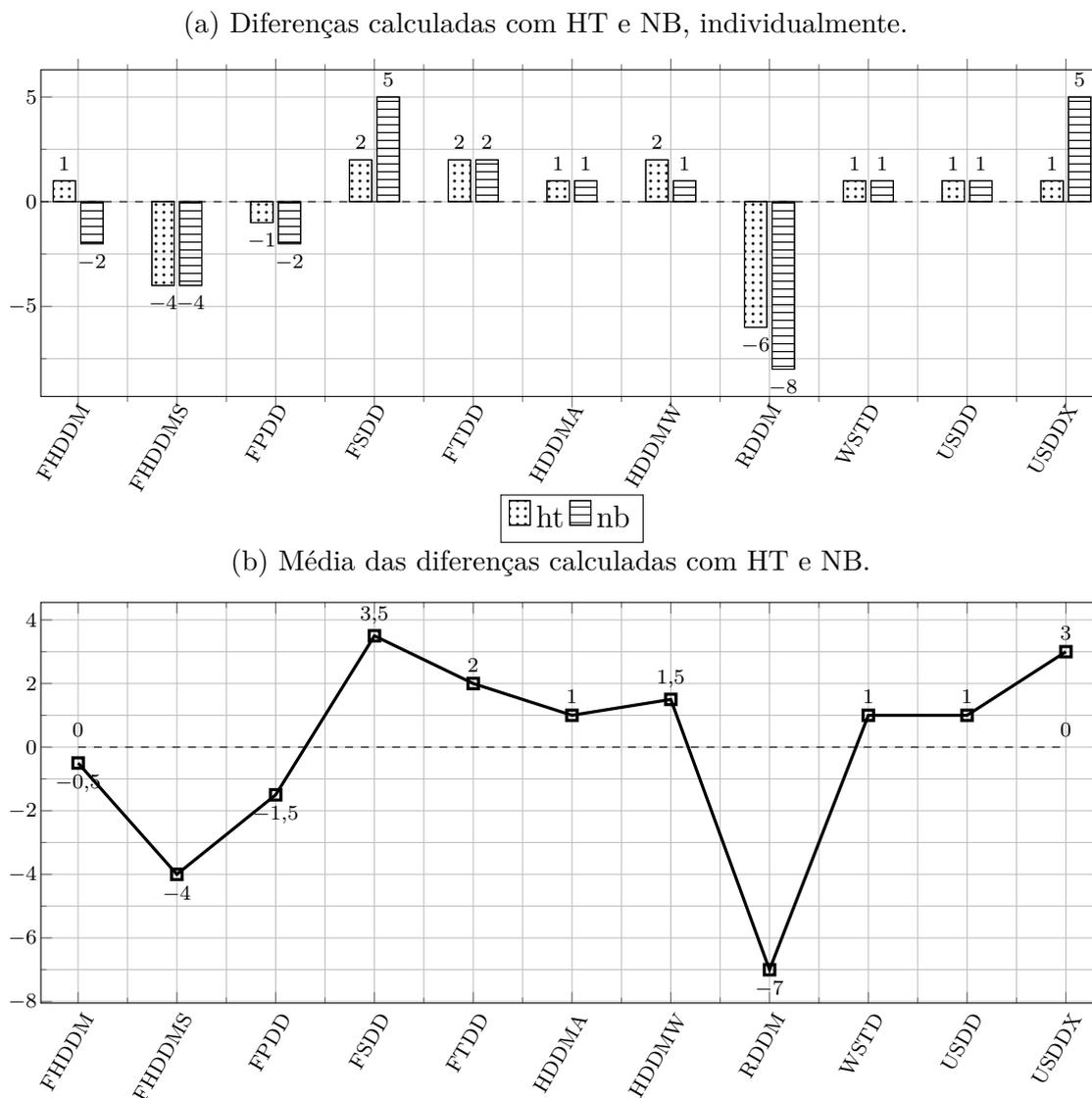
Fonte: O AUTOR (2020).

Tabela 13 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _g | HDDM _w | RDDM | WSTD | USDD | USDDX |
|--------------|--------------|--------------|----------|-----------------|----------|-------------------|-------------------|--------------|---------------|----------------|----------------|
| A-20k-Agr | 116,51 | 121,04 | 67,39 | 94,07 | 66,48 | 87,95 | 132,61 | 52,64 | 140,66 | 58,73 | 137,40 |
| A-20k-LED | 44,97 | 57,04 | 0,46 | 0,81 | 0,45 | 19,59 | 93,59 | 15,06 | 11,99 | 25,83 | 25,83 |
| A-20k-Mix | 306,08 | 306,08 | 345,14 | 345,14 | 345,14 | 459,62 | 426,00 | 183,33 | 374,36 | 118,93 | 538,22 |
| A-20k-RBF | 25,85 | 39,94 | 3,57 | 10,34 | 3,62 | 20,44 | 36,93 | 19,67 | 10,96 | 12,14 | 11,66 |
| A-20k-Sine | 284,58 | 284,58 | 283,03 | 293,51 | 288,32 | 221,29 | 363,52 | 122,49 | 299,98 | 109,74 | 294,48 |
| A-20k-Wav | 51,84 | 88,85 | 7,09 | 34,64 | 22,15 | 27,17 | 114,05 | 25,68 | 35,21 | 11,02 | 35,14 |
| A-130k-Agr | 531,93 | 573,75 | 585,69 | 721,14 | 594,73 | 289,91 | 776,10 | 142,37 | 835,10 | 348,79 | 667,67 |
| A-130k-LED | 59,68 | 69,91 | 4,83 | 26,87 | 4,86 | 87,89 | 54,67 | 49,55 | 9,86 | 164,68 | 164,68 |
| A-130k-Mix | 2037,77 | 2037,77 | 2277,70 | 2277,70 | 2277,70 | 1375,12 | 2836,13 | 514,13 | 2468,60 | 770,44 | 3461,67 |
| A-130k-RBF | 89,45 | 133,69 | 19,37 | 136,42 | 56,55 | 89,95 | 98,74 | 55,05 | 115,92 | 76,07 | 67,88 |
| A-130k-Sine | 1894,08 | 1871,53 | 1919,41 | 1987,90 | 1953,63 | 642,88 | 2453,78 | 252,86 | 1992,14 | 730,59 | 1963,27 |
| A-130k-Wav | 447,25 | 406,65 | 166,23 | 496,92 | 367,31 | 199,48 | 414,47 | 77,73 | 442,11 | 308,52 | 302,39 |
| A-890k-Agr | 4038,41 | 4215,80 | 4124,63 | 4630,12 | 4142,22 | 765,94 | 5061,79 | 210,27 | 4777,01 | 2302,68 | 3682,83 |
| A-890k-LED | 76,17 | 87,26 | 125,22 | 218,26 | 112,85 | 339,96 | 59,45 | 75,76 | 16,19 | 629,60 | 629,60 |
| A-890k-Mix | 13633,88 | 13633,88 | 14959,61 | 14959,61 | 14959,61 | 4215,07 | 18776,21 | 940,79 | 16327,63 | 5254,98 | 17604,22 |
| A-890k-RBF | 227,29 | 151,45 | 181,23 | 465,73 | 323,00 | 211,93 | 112,91 | 51,81 | 168,54 | 489,86 | 380,52 |
| A-890k-Sine | 11627,48 | 9440,65 | 12156,83 | 12724,30 | 12527,86 | 1368,98 | 12427,04 | 285,46 | 11163,95 | 5005,57 | 10217,06 |
| A-890k-Wav | 1070,72 | 610,60 | 1633,53 | 1859,05 | 2051,98 | 695,58 | 520,56 | 86,27 | 1160,33 | 2153,55 | 1348,07 |
| G-20k-Agr | 13,21 | 7,71 | 4,01 | 5,48 | 3,87 | 13,13 | 12,04 | 11,84 | 3,86 | 12,16 | 6,58 |
| G-20k-LED | 6,21 | 13,33 | 0,29 | 0,65 | 0,28 | 8,91 | 7,49 | 11,76 | 0,37 | 8,92 | 8,92 |
| G-20k-Mix | 1,06 | 0,84 | 3,48 | 3,53 | 3,48 | 0,88 | 0,97 | 8,14 | 5,24 | 1,15 | 2,49 |
| G-20k-RBF | 4,94 | 4,31 | 2,61 | 3,58 | 2,68 | 9,94 | 4,85 | 11,83 | 1,17 | 4,71 | 3,95 |
| G-20k-Sine | 1,45 | 0,90 | 3,19 | 2,30 | 3,72 | 0,89 | 0,83 | 3,57 | 5,58 | 4,26 | 1,21 |
| G-20k-Wav | 0,45 | 0,64 | 0,31 | 0,30 | 0,30 | 0,49 | 1,42 | 1,43 | 0,28 | 0,31 | 0,48 |
| G-130k-Agr | 76,86 | 27,87 | 59,45 | 65,85 | 57,85 | 132,99 | 67,27 | 129,43 | 42,13 | 71,50 | 67,02 |
| G-130k-LED | 13,38 | 16,97 | 1,28 | 9,89 | 1,19 | 71,71 | 9,36 | 51,60 | 1,35 | 54,24 | 54,24 |
| G-130k-Mix | 0,88 | 12,22 | 28,83 | 33,44 | 30,32 | 99,87 | 1,06 | 68,91 | 20,83 | 37,90 | 76,28 |
| G-130k-RBF | 17,44 | 14,96 | 14,64 | 24,40 | 24,94 | 51,30 | 11,40 | 46,85 | 4,00 | 32,40 | 25,83 |
| G-130k-Sine | 1,08 | 0,99 | 16,64 | 8,57 | 12,23 | 41,45 | 12,28 | 65,55 | 15,64 | 123,25 | 1,68 |
| G-130k-Wav | 64,74 | 33,03 | 21,33 | 76,30 | 48,06 | 62,82 | 22,11 | 64,89 | 17,99 | 93,00 | 43,83 |
| G-890k-Agr | 411,57 | 66,77 | 453,48 | 511,37 | 453,50 | 464,38 | 409,72 | 71,56 | 292,95 | 505,18 | 715,76 |
| G-890k-LED | 17,73 | 22,15 | 48,37 | 67,03 | 49,08 | 273,23 | 10,48 | 59,50 | 1,95 | 153,15 | 153,15 |
| G-890k-Mix | 0,81 | 0,83 | 247,61 | 223,79 | 220,70 | 1062,98 | 1,06 | 92,94 | 394,77 | 636,14 | 262,28 |
| G-890k-RBF | 30,89 | 9,41 | 128,12 | 131,11 | 162,65 | 193,52 | 9,97 | 43,46 | 15,40 | 210,05 | 180,72 |
| G-890k-Sine | 1,39 | 1,08 | 207,09 | 12,11 | 178,08 | 409,75 | 1,09 | 95,84 | 312,75 | 627,17 | 88,70 |
| G-890k-Wav | 124,52 | 10,88 | 421,03 | 137,51 | 403,37 | 365,24 | 31,74 | 42,07 | 6,30 | 849,90 | 513,81 |
| Vitórias | 1 | 0 | 1 | 5 | 4 | 1 | 1 | 0 | 1 | 1 | 5 |
| Derrotas | 3 | 4 | 3 | 0 | 2 | 0 | 0 | 8 | 0 | 0 | 0 |
| Média Postos | 6,33 | 6,97 | 7,22 | 5,19 | 6,56 | 5,22 | 5,69 | 6,94 | 6,36 | 5,00 | 4,50 |

Fonte: O AUTOR (2020).

Figura 11 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise.



Fonte: O AUTOR (2020).

As Tabelas 14 e 15 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o método USDD obteve os melhores resultados com HT (2,97) e NB (3,33).

A Figura 12 (a) apresenta o resumo dos resultados com HT e NB apresentados no apêndice E nas Tabelas 96 e 97. Em termos das hipóteses testadas, o USDD forneceu o maior número de superioridades estatísticas com HT e com NB, ambos com 5 vitórias e nenhuma derrota. Entretanto, apenas os métodos FHDDM, HDDMA, HDDMW, RDDM e USDDX foram inferiores estatisticamente ao USDD com o HT e NB. O método com maior número de derrotas usando HT (10) e NB (10) foi o RDDM, logo, tratando-se da

ineficiência, ele foi, em termos absolutos e em termos estatísticos, o pior método avaliado. Vale ressaltar que o RDDM não obteve nenhuma superioridade estatística em relação aos demais métodos avaliados, pois todos os demais métodos foram superiores estatisticamente a ele com HT e NB.

É fácil notar que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos métodos USDD, FHDDMS, FHDDM, WSTD, FPDD, FSDD e FTDD, todos com avaliação acima de zero, mostrados na Figura 12 (b). Apesar disso, vale a pena destacar que o desempenho do método USDD foi sólido na maioria das situações e líder em vitórias com HT e NB. A Figura 12 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que os métodos USDD e FHDDMS obtiveram as maiores médias no desempenho acumulado, seguidos pelos métodos FHDDM (2), WSTD (2), FPDD, FSDD e FTDD com 1,5, USDDX com 0,5 e os demais métodos ficaram com média negativa em número de vitórias, ou seja, foram inferiores mais vezes que superiores estatisticamente.

Tabela 14 – Médias dos MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------------|-------|-------------|-------------|-------------|
| A-20k-Agr | 0,88 | 0,93 | 0,79 | 0,83 | 0,79 | 0,81 | 0,86 | 0,73 | 0,93 | 0,84 | 0,88 |
| A-20k-LED | 0,23 | 0,32 | 0,03 | 0,03 | 0,02 | 0,13 | 0,37 | 0,03 | 0,12 | 0,38 | 0,38 |
| A-20k-Mix | 1,00 | 1,00 | 0,95 | 0,95 | 0,95 | 0,97 | 0,99 | 0,86 | 0,97 | 1,00 | 0,90 |
| A-20k-RBF | 0,62 | 0,66 | 0,56 | 0,66 | 0,61 | 0,47 | 0,61 | 0,36 | 0,72 | 0,66 | 0,65 |
| A-20k-Sine | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 1,00 | 1,00 | 0,85 | 0,98 | 1,00 | 0,97 |
| A-20k-Wav | 0,48 | 0,48 | 0,54 | 0,52 | 0,56 | 0,39 | 0,47 | 0,32 | 0,51 | 0,63 | 0,39 |
| A-130k-Agr | 0,80 | 0,86 | 0,93 | 0,94 | 0,93 | 0,87 | 0,70 | 0,22 | 0,94 | 0,77 | 0,79 |
| A-130k-LED | 0,11 | 0,15 | 0,02 | 0,01 | 0,02 | 0,07 | 0,16 | 0,01 | 0,05 | 0,17 | 0,17 |
| A-130k-Mix | 1,00 | 1,00 | 0,94 | 0,94 | 0,94 | 0,97 | 1,00 | 0,82 | 0,95 | 1,00 | 0,86 |
| A-130k-RBF | 0,45 | 0,46 | 0,68 | 0,70 | 0,69 | 0,53 | 0,34 | 0,00 | 0,66 | 0,52 | 0,51 |
| A-130k-Sine | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,96 | 0,99 | 0,73 | 0,98 | 1,00 | 0,96 |
| A-130k-Wav | 0,38 | 0,34 | 0,56 | 0,44 | 0,53 | 0,19 | 0,26 | 0,12 | 0,44 | 0,65 | 0,40 |
| A-890k-Agr | 0,84 | 0,85 | 0,98 | 0,97 | 0,98 | 0,69 | 0,69 | 0,13 | 0,94 | 0,82 | 0,81 |
| A-890k-LED | 0,05 | 0,05 | 0,01 | 0,01 | 0,00 | 0,03 | 0,07 | 0,01 | 0,03 | 0,06 | 0,06 |
| A-890k-Mix | 1,00 | 1,00 | 0,94 | 0,94 | 0,94 | 0,95 | 1,00 | 0,40 | 0,95 | 1,00 | 0,83 |
| A-890k-RBF | 0,33 | 0,26 | 0,72 | 0,70 | 0,74 | 0,42 | 0,15 | 0,05 | 0,49 | 0,52 | 0,51 |
| A-890k-Sine | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,97 | 0,99 | 0,37 | 0,97 | 1,00 | 0,95 |
| A-890k-Wav | 0,25 | 0,14 | 0,54 | 0,38 | 0,50 | 0,26 | 0,12 | 0,05 | 0,26 | 0,61 | 0,32 |
| Vitórias | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 5 | 1 |
| Derrotas | 1 | 0 | 0 | 0 | 0 | 4 | 3 | 10 | 0 | 0 | 1 |
| Média Postos | 5,00 | 4,67 | 5,94 | 5,61 | 6,00 | 7,00 | 5,97 | 10,89 | 5,11 | 2,97 | 6,83 |

Fonte: O AUTOR (2020).

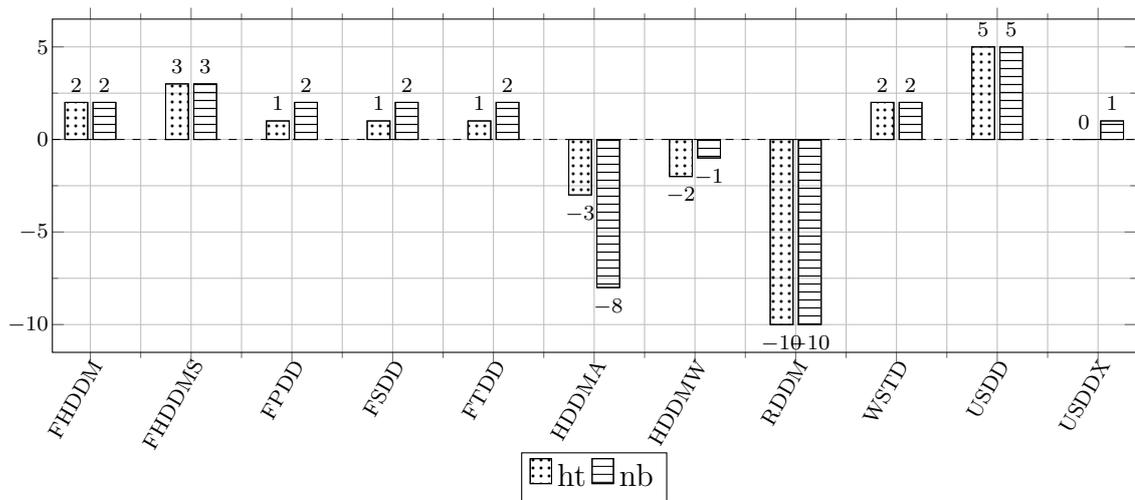
Tabela 15 – Médias dos MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------------|-------|-------------|-------------|-------------|
| A-20k-Agr | 0,91 | 0,93 | 0,78 | 0,87 | 0,77 | 0,84 | 0,86 | 0,71 | 0,96 | 0,89 | 0,90 |
| A-20k-LED | 0,49 | 0,59 | 0,08 | 0,10 | 0,07 | 0,17 | 0,52 | 0,01 | 0,24 | 0,47 | 0,47 |
| A-20k-Mix | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 1,00 | 1,00 | 0,97 | 1,00 | 1,00 | 1,00 |
| A-20k-RBF | 0,68 | 0,66 | 0,65 | 0,71 | 0,69 | 0,55 | 0,60 | 0,39 | 0,74 | 0,68 | 0,66 |
| A-20k-Sine | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 0,99 | 0,99 | 0,91 | 0,99 | 1,00 | 0,98 |
| A-20k-Wav | 0,71 | 0,85 | 0,47 | 0,65 | 0,57 | 0,48 | 0,85 | 0,33 | 0,67 | 0,62 | 0,70 |
| A-130k-Agr | 0,62 | 0,61 | 0,85 | 0,90 | 0,85 | 0,58 | 0,55 | 0,26 | 0,86 | 0,71 | 0,68 |
| A-130k-LED | 0,21 | 0,26 | 0,08 | 0,17 | 0,08 | 0,11 | 0,21 | 0,01 | 0,18 | 0,43 | 0,43 |
| A-130k-Mix | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 0,95 | 1,00 | 0,90 | 1,00 | 1,00 | 0,96 |
| A-130k-RBF | 0,34 | 0,34 | 0,65 | 0,58 | 0,61 | 0,26 | 0,29 | 0,06 | 0,56 | 0,37 | 0,36 |
| A-130k-Sine | 0,99 | 0,97 | 0,99 | 0,98 | 0,99 | 0,72 | 0,95 | 0,52 | 0,97 | 1,00 | 0,96 |
| A-130k-Wav | 0,66 | 0,67 | 0,53 | 0,72 | 0,63 | 0,29 | 0,57 | 0,07 | 0,68 | 0,73 | 0,68 |
| A-890k-Agr | 0,29 | 0,28 | 0,84 | 0,83 | 0,84 | 0,26 | 0,24 | 0,15 | 0,55 | 0,37 | 0,36 |
| A-890k-LED | 0,10 | 0,11 | 0,05 | 0,12 | 0,05 | 0,05 | 0,09 | 0,01 | 0,08 | 0,28 | 0,28 |
| A-890k-Mix | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,59 | 1,00 | 0,31 | 0,99 | 1,00 | 0,80 |
| A-890k-RBF | 0,14 | 0,13 | 0,47 | 0,36 | 0,46 | 0,12 | 0,11 | 0,03 | 0,24 | 0,15 | 0,15 |
| A-890k-Sine | 0,92 | 0,68 | 0,94 | 0,92 | 0,94 | 0,40 | 0,61 | 0,20 | 0,78 | 1,00 | 0,85 |
| A-890k-Wav | 0,34 | 0,29 | 0,52 | 0,47 | 0,57 | 0,12 | 0,21 | 0,04 | 0,39 | 0,71 | 0,57 |
| Vitórias | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 0 | 2 | 5 | 2 |
| Derrotas | 1 | 0 | 0 | 0 | 0 | 9 | 3 | 10 | 0 | 0 | 1 |
| Média Postos | 4,69 | 4,86 | 6,00 | 4,83 | 5,67 | 8,83 | 6,64 | 11,00 | 4,67 | 3,33 | 5,47 |

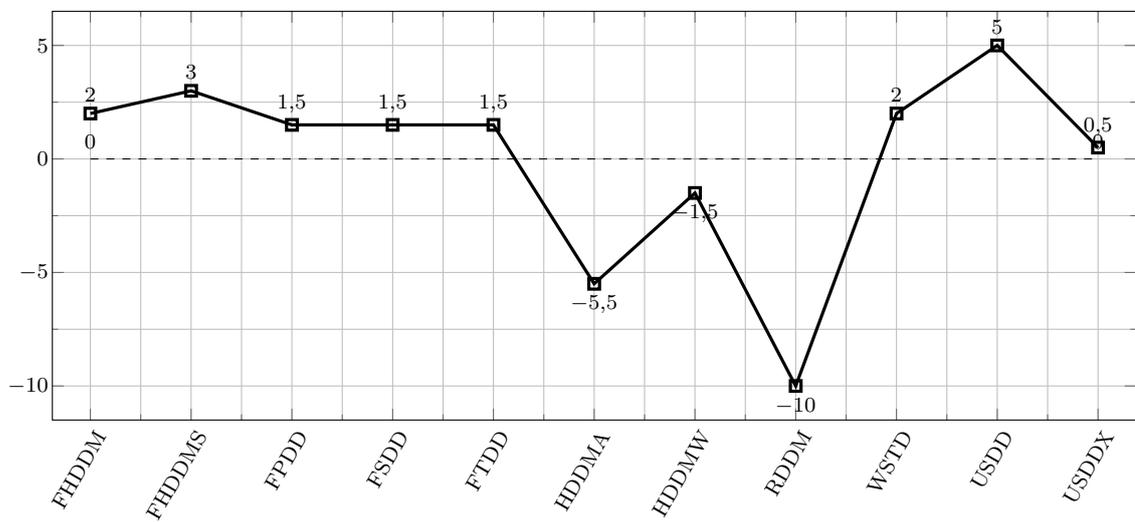
Fonte: O AUTOR (2020).

Figura 12 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



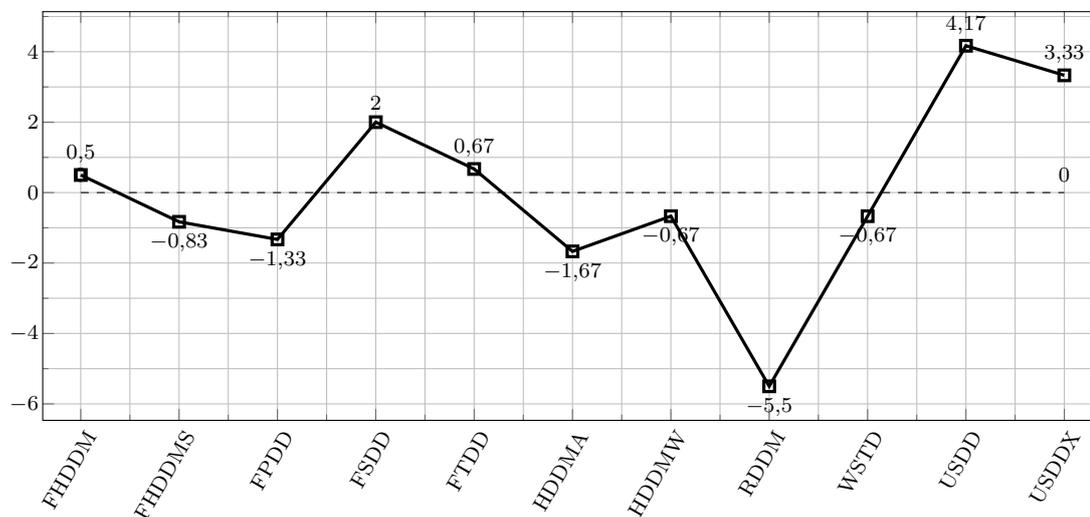
Fonte: O AUTOR (2020).

A Figura 13 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 13 (a), é feito o rank da média das diferenças. A Figura 13 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 13 (a) ou no da Figura 13 (b), o detector de maior valor absoluto foi o USDD com média 4,17 e 9,25, respectivamente. Vale ressaltar que, por meio do rank da média (Figura 13 (a)), é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários. O RDDM, apesar de ser um excelente detector em acurácia, nesses cenários e critérios de avaliação acabou apresentando o pior resultado nessa metodologia de avaliação.

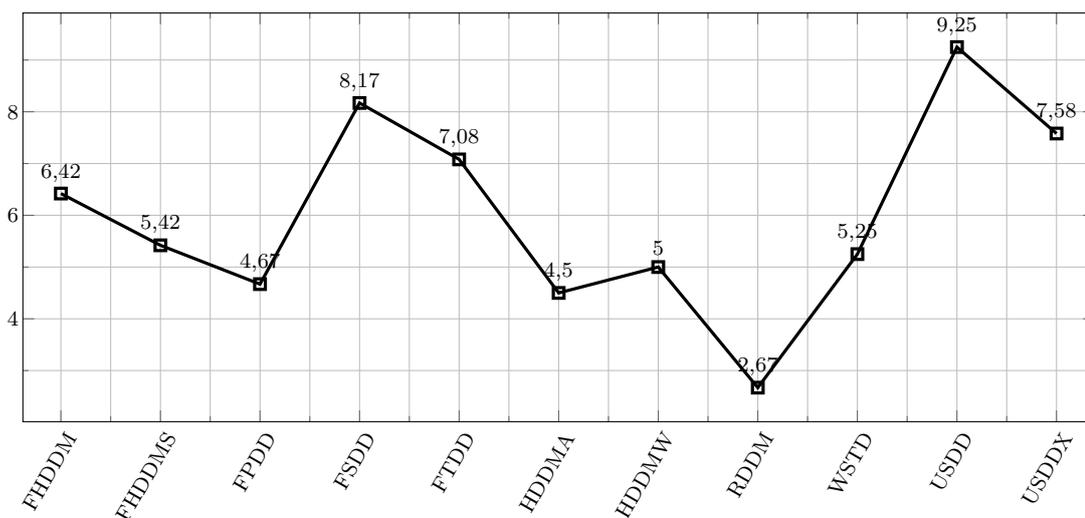
O USDD foi o único método no qual em todas as avaliações não obteve derrota. Além disso, ele também foi o método de melhor classificação ordenada por cálculo de média dos ranks e por cálculo de média dos postos. Assim, o USDD apresenta-se como um novo método competitivo para DMC. Vale ressaltar que o USDDX foi o segundo melhor método classificado e obteve três (3) derrotas que foi justamente para o USDD, uma na avaliação da acurácia com HT e outras duas na avaliação do MCC, uma com o HT e outra com o NB. Em termos de acurácia da classificação, o USDD obteve os melhores resultados com o HT, enquanto que o USDDX obteve os melhores resultados com o NB.

Figura 13 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com os métodos propostos de única análise.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.



Fonte: O AUTOR (2020).

As Figuras 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 e 24 apresentam a correlação de eficiência entre os critérios de avaliação e desempenho entre métodos para cada método avaliado nesta Seção 5. A eficiência é calculada com base em vitórias (superioridades estatísticas) obtidas em cada critério de avaliação e são apresentados em termos percentuais (%) individualmente em escala de 0 a 1 (100%). Primeiramente, foram somadas as vitórias nos critérios de acurácia, MTR, MCC e então foram calculados os percentuais de cada componente com base no total de vitórias do método. Apesar de as figuras considerarem as vitórias individuais, não se deseja identificar o maior ou menor número de vitórias, contudo, pretende-se discutir a correlação de desempenho (ou seja, o

viés) de cada método nos critérios de avaliação adotados. Essa análise permite mapear os pontos fortes e fracos dos métodos na avaliação de desempenho.

Particularmente, a tarefa do método DMC em uma AADE é induzida pelo classificador. Assim sendo, uma diminuição da taxa de FP da classificação contribui para melhoria da acurácia da classificação e, conseqüentemente, pode contribuir para DMC, quando uma MC ocorrer. Entretanto, um método de DMC tende a ter comportamento diferente ao mudar o algoritmo de classificação, mas geralmente mantém o viés dos resultados em termos de acurácia da classificação e DMC.

Devido ao fato do objetivo da avaliação de desempenho ser três competências, o viés analisado do detector é em relação a elas. Em contraste com o viés, quando a variância do detector é alta, ele tende a sofrer mais com ruídos ou dados específicos. Assim, para os detectores, o ideal é ter um método com viés elevado para detecção, mas que também não tenha uma variância muito alta para não sofrer com ruídos e dados aleatórios.

Por exemplo, na Figura 14, o FHDDM apresentou seu melhor desempenho na competência MCC com 50% do total em ambos os classificadores HT e NB. Note-se que houve um equilíbrio no desempenho para ambos os classificadores, repetindo resultados iguais nos três critérios. Por fim, o critério MTR, com 17%, apresentou o menor percentual de desempenho entre os critérios de avaliação, logo, o MTR não foi o ponto forte do método FHDDM.

Na Figura 19, o HDDMA apresentou o mesmo desempenho nas três competências com 33% do total (houve um arredondamento de 33,33 para 33%) em ambos os classificadores HT e NB. Note-se que houve um equilíbrio no desempenho do método para ambos os classificadores, repetindo resultados iguais. Por fim, o HDDMA apresentou um equilíbrio em termos percentuais entre os três critérios, mostrando-se ser um método versátil e de difícil análise dos pontos fortes e fracos. O equilíbrio citado se trata da semelhança entre o desempenho nas métricas avaliadas.

Na Figura 17, o FSDD apresentou seu melhor desempenho na competência MTR, com 67% e 50% nos classificadores HT e NB, respectivamente. Note-se que em termos de acurácia o aproveitamento foi melhor com NB. Por fim, a acurácia com 0% com HT apresentou o menor percentual de desempenho entre os critérios de avaliação, logo, a acurácia com HT não foi o ponto forte do método FSDD.

Na Figura 21, o RDDM apresentou seu melhor desempenho na competência acurácia com 100% no NB. Note-se que em termos de acurácia o aproveitamento foi melhor com NB, entretanto o desempenho com o HT não se repetiu. O ponto forte do RDDM foi a acurácia com NB e como ponto fraco estão os critérios de MTR e MCC. Vale ressaltar que o desempenho apresentado pelo RDDM colabora com o entendimento que ele seja um bom método para ser utilizado com o NB para melhoria da acurácia.

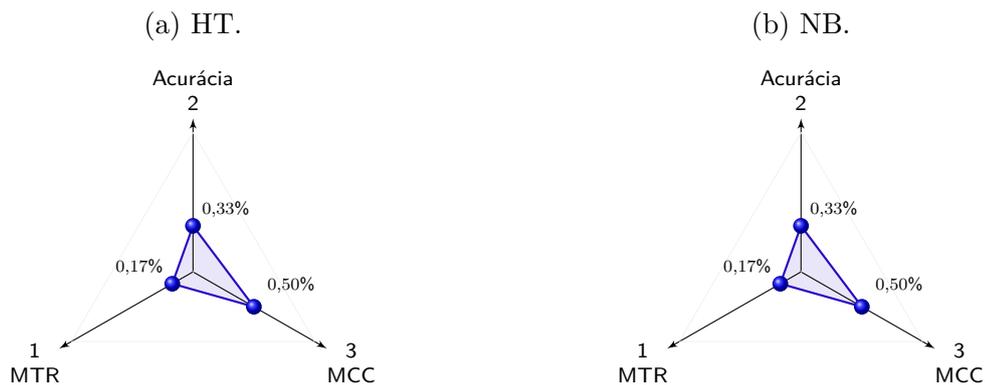
Na Figura 23, o USDD apresentou seu melhor desempenho na competência acurácia com 54% e 50% nos classificadores HT e NB, respectivamente. Note-se que em termos de

acurácia o aproveitamento foi melhor com HT. Por fim, o MTR com 8% com HT e NB apresentou o menor percentual de desempenho entre os critério de avaliação, logo, o MTR não foi o ponto forte do método USDD.

Na Figura 24, o USDDX apresentou seu melhor desempenho na competência também na acurácia com 75% e 53% nos classificadores HT e NB, respectivamente. Note-se que em termos de acurácia o aproveitamento também foi melhor com HT, levando em consideração a correlação entre os critérios. Por fim, o MCC com 13% com HT e NB apresentou o menor percentual de desempenho entre os critérios de avaliação, logo, o MCC não foi o ponto forte do método USDDX.

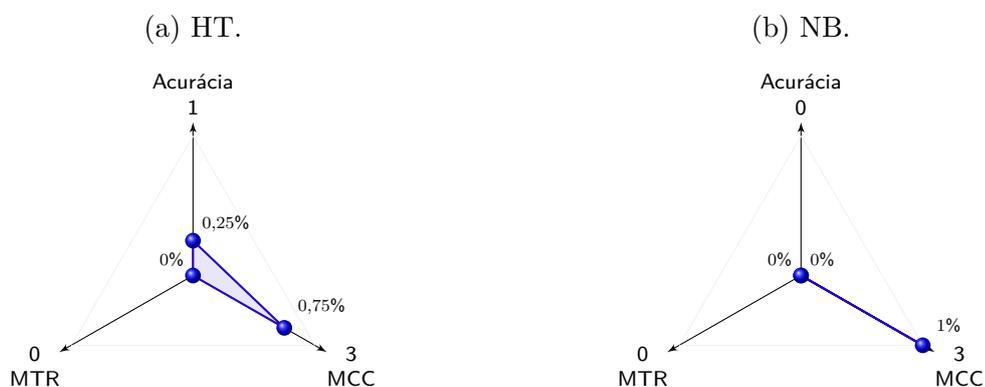
Por fim, tratando-se do objetivo a ser otimizado na tarefa a ser realizada, pode-se optar por priorizar o uso dos métodos de melhor desempenho e com pontos fortes em tal critério. Essa avaliação apresentou o desempenho e metodologia para avaliar pontos fortes e fracos entre métodos avaliados com base em técnicas estatísticas e correlação percentual dos critérios avaliados.

Figura 14 – Correlação de eficiência do FHDDM com HT e NB.



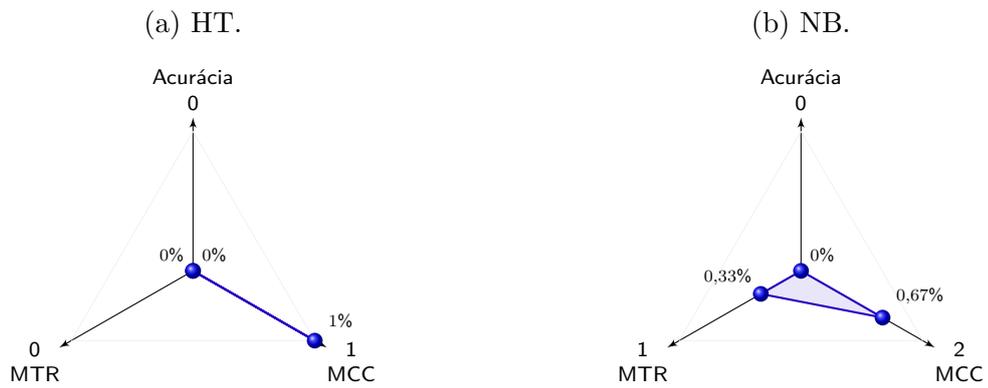
Fonte: O AUTOR (2020).

Figura 15 – Correlação de eficiência do FHDDMS com HT e NB.



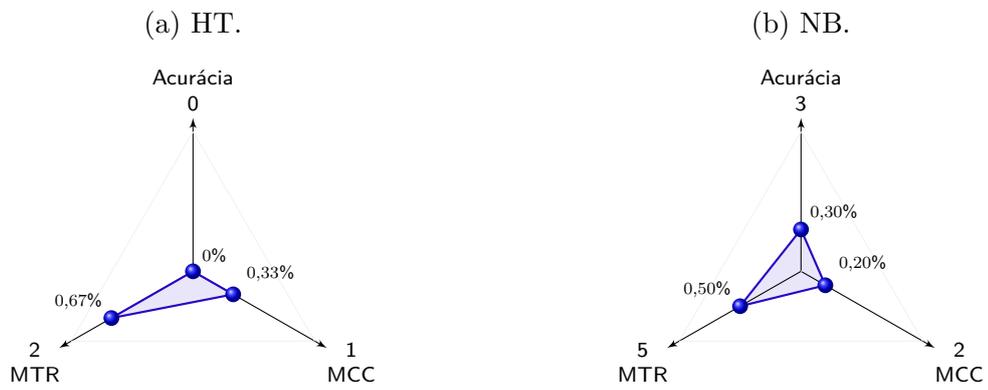
Fonte: O AUTOR (2020).

Figura 16 – Correlação de eficiência do FPDD com HT e NB.



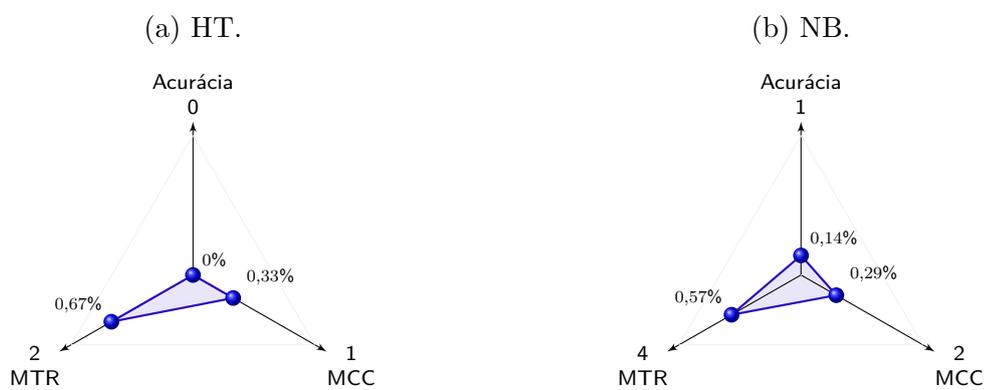
Fonte: O AUTOR (2020).

Figura 17 – Correlação de eficiência do FSDD com HT e NB.



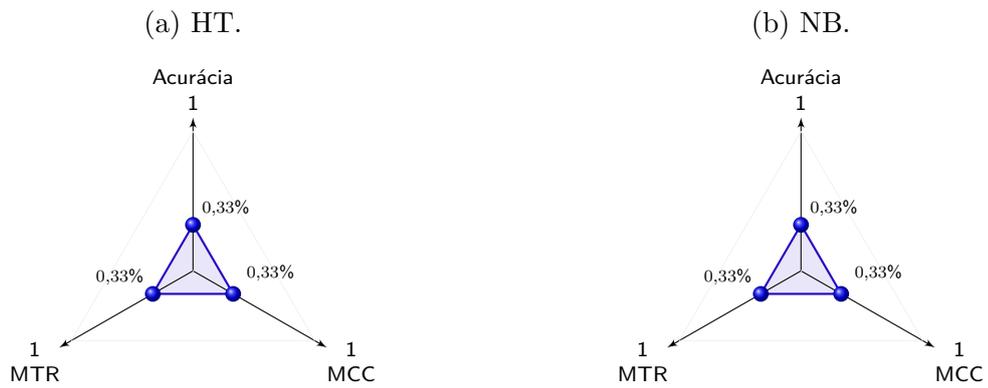
Fonte: O AUTOR (2020).

Figura 18 – Correlação de eficiência do FTDD com HT e NB.



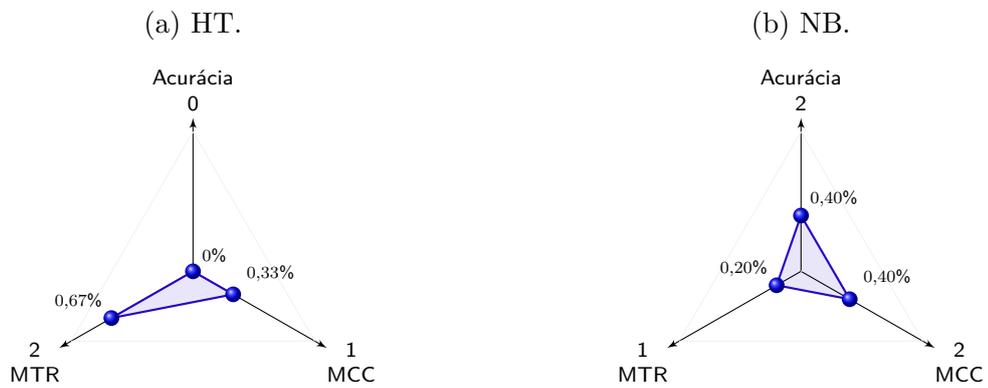
Fonte: O AUTOR (2020).

Figura 19 – Correlação de eficiência do HDDDMA com HT e NB.



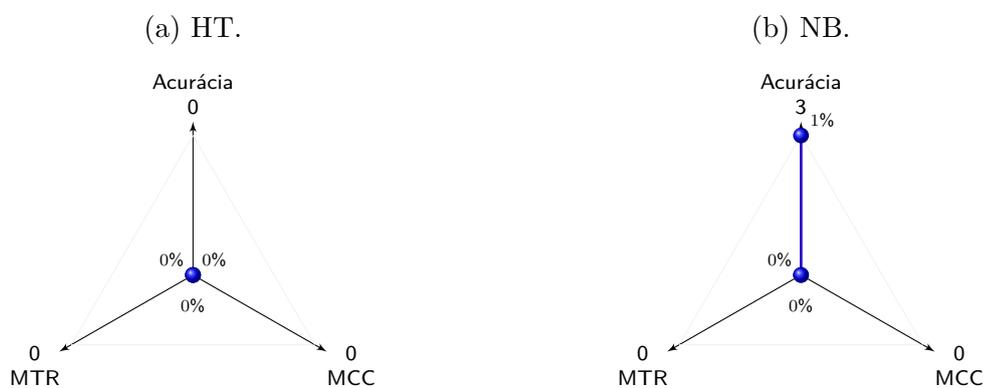
Fonte: O AUTOR (2020).

Figura 20 – Correlação de eficiência do HDDMW com HT e NB.



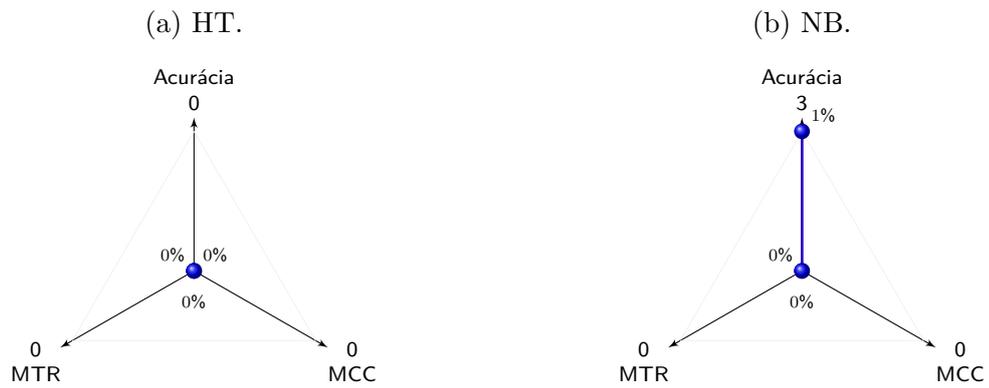
Fonte: O AUTOR (2020).

Figura 21 – Correlação de eficiência do RDDM com HT e NB.



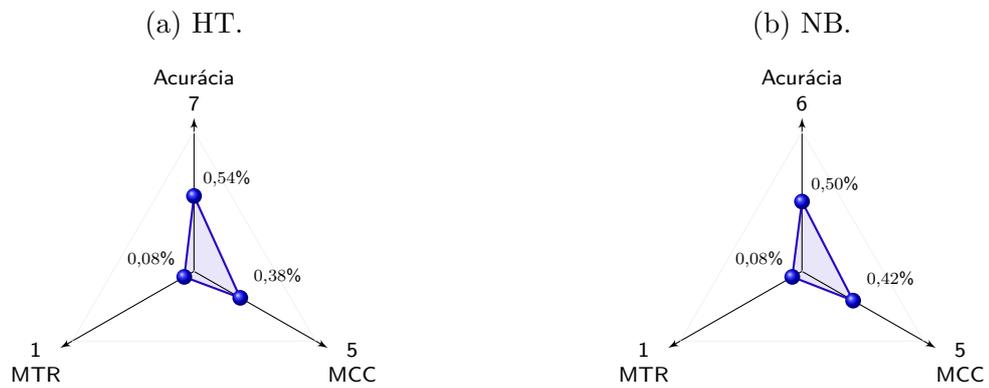
Fonte: O AUTOR (2020).

Figura 22 – Correlação de eficiência do WSTD com HT e NB.



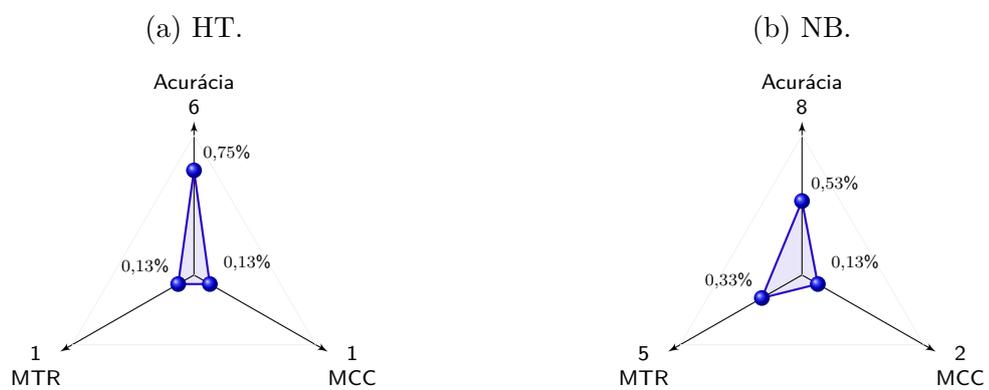
Fonte: O AUTOR (2020).

Figura 23 – Correlação de eficiência do USDD com HT e NB.



Fonte: O AUTOR (2020).

Figura 24 – Correlação de eficiência do USDDX com HT e NB.



Fonte: O AUTOR (2020).

5.4 CONSIDERAÇÕES FINAIS

Esta seção apresentou os métodos USDD e USDDX que são baseados em um algoritmo novo, eficiente e, finalmente, simples de DMC. O critério de decisão por trás dos dois métodos é consideravelmente simples e utiliza uma janela deslizante para calcular a estatística do teste, baseada em proporção. Apesar da simplicidade do algoritmo, ele tem-se mostrado eficiente, principalmente em casos com conceitos maiores e mudanças de conceito graduais.

Os métodos USDD e USDDX foram testados com dois algoritmos de aprendizado diferentes, a saber: *Hoeffding Tree* (HT) e *Naive Bayes* (NB). Foram considerados na avaliação 36 diferentes parametrizações de geradores de artificiais, provenientes de 6 geradores de dados e sete conjuntos de dados do mundo real, contra nove dos melhores detectores de MC disponíveis e avaliados no apêndice B. Os resultados de precisão dos experimentos também foram avaliados estatisticamente usando o teste de *Wilcoxon Signed-Rank*, juntamente com três critérios de avaliação que consideram avaliar competências diferentes que se espera que um detector ofereça bons resultados. Por fim, empiricamente, foi observado que conceitos mais longos levam a modelos mais precisos e, nessas situações, o USDD pode fornecer resultados melhores do que com tamanhos de janela menores.

6 MOAMANAGER

Esta seção apresenta a MOAManager (MACIEL; SANTOS; BARROS, 2020), uma estrutura com interface web para realizar gestão de experimentos com fluxos de dados que exploram os recursos da estrutura do MOA. A ferramenta pode ser vista como uma extensão ao MOA, pois, para que ela funcione, é preciso adicionar um complemento ao código fonte do MOA. Ao realizar manualmente experimentos com fluxos de dados no MOA, é exigido um grande esforço, podendo resultar em custo excessivo de tempo para conclusão das experimentações. A proposta da ferramenta MOAManager é facilitar e otimizar as etapas de experimentações e análise de dados, integrando-se à estrutura do MOA por meio da exportação dos resultados.

Trabalhar em ambientes de fluxo de dados (GAMA, 2010; GONÇALVES JR. et al., 2014) geralmente envolve lidar com grandes quantidades de informações. Há várias etapas nesse processo, incluindo preparação, execução e análise de dados. Normalmente, esses são estágios que exigem tempo considerável e são bastante suscetíveis a erros humanos quando executados manualmente. Portanto, a adoção de uma ferramenta de automação que auxilie nessas atividades deve resultar em ganhos consideráveis de produtividade e qualidade.

O MOAManager é uma ferramenta de interface *web*, amigável, simples e fácil de usar que fornece vários recursos que foram projetados para oferecer suporte a grandes experimentos usando fluxos de dados. Em outras palavras, ela pode ser vista como uma camada adicional inserida na estrutura do MOA (BIFET et al., 2010) - uma estrutura popular para executar experimentos de fluxos de dados - tornando o processamento final mais transparente. Além disso, ela é de código aberto, está disponível gratuitamente e um resumo das informações relevantes para instalação e uso são apresentadas abaixo.

- **Versão do software:** v1.1;
- **Repositório:** <<https://github.com/brunom4ciel/moamanager>>;
- **Licença de Software:** GNU General Public License versão 2 ou posterior;
- **Sistema Operacional:** Plataforma independente, Ubuntu 18;
- **Manual do usuário:** <<https://github.com/brunom4ciel/moamanager/wiki/use-r-guide>>;
- **Manual do desenvolvedor:** <<https://github.com/brunom4ciel/moamanager/wiki>>.

Vários recursos estão disponíveis, entre os quais destacam-se:

1. Interface amigável para geração automática de *Script* MOA com suporte a diferentes parametrizações;
2. Possibilidade de multiprocessamento nos experimentos, permitindo fácil controle do uso simultâneo de recursos computacionais multi-core;
3. Retomada da execução em caso de interrupções, voluntárias ou não;
4. Possibilidade de receber notificação por e-mail ao final dos experimentos desde que a máquina esteja conectada à internet;
5. Extração e agregação de múltiplos resultados de acurácia, memória ou tempo de execução, a partir de um grupo de arquivos processados, com ou sem seus respectivos intervalos de confiança e com suporte à sua exportação para diferentes formatos (e.g. *HyperText Markup Language* (HTML), \LaTeX , *Comma-Separated Values* (CSV), etc);
6. Cálculo de métricas relacionadas a métodos de detecção de mudança de conceito (GONÇALVES JR. et al., 2014; BARROS; SANTOS, 2018; BARROS et al., 2017; BARROS; HIDALGO; CABRAL, 2018), por exemplo: a matriz de confusão (falsos e verdadeiros positivos e negativos), acurácia e *recall* (PERRY; KENT; BERRY, 1955), F-Measure (F1 score), MCC (MATTHEWS, 1975), etc .; e
7. Cálculo de métricas estatísticas para análise de dados, a exemplo de: intervalo de confiança, teste de Friedman (FRIEDMAN, 1937), *Bonferroni-Dunn*, *Nemenyi*, *Shaffer*, *Holm*, *Bergmann-Hommel*, *Kullback-Leibler*, *Minkowski*, etc.

A MOAManager está sendo usado pelo grupo de pesquisa sob supervisão do Prof. Roberto Souto Maior de Barros do Centro de Informática da Universidade Federal de Pernambuco há mais de quatro anos. O rastreamento interno do grupo sugere que, quando foi desenvolvido, não havia nenhuma outra ferramenta disponível para apoiar os usuários e ajudá-los a superar as dificuldades e limitações presentes no MOA e suas interfaces com relação à automação de tarefas para executar grandes experimentos com vários métodos e conjuntos de dados relacionados para fluxos de dados.

O restante da seção está organizada nos seguintes tópicos: 6.1 descreve em detalhes a maioria dos recursos da ferramenta MOAManager; 6.2 apresenta alguns detalhes de implementação; 6.3 compartilha a experiência do grupo no uso e desenvolvimento da ferramenta desde 2015, e 6.4 fornece as conclusões.

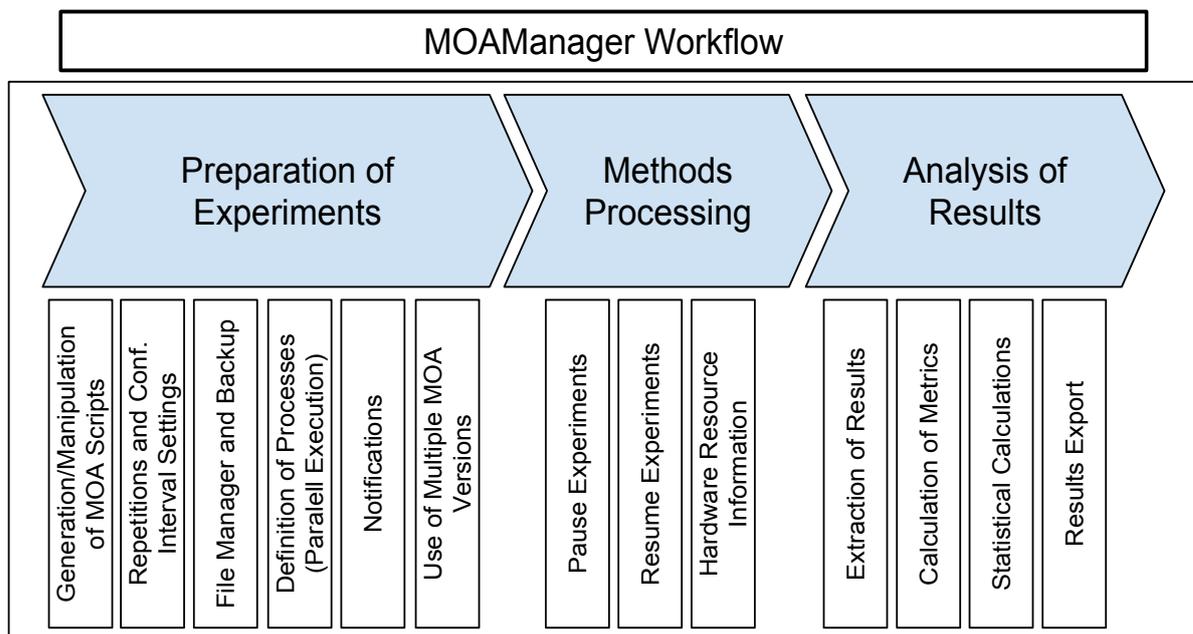
6.1 PRINCIPAIS FUNCIONALIDADES

Esta seção descreve os principais recursos da ferramenta MOAManager. Para uma melhor organização, a apresentação é dividida em três fases. Na primeira fase, ocorre a

preparação dos dados, onde o usuário especifica as parametrizações dos *scripts*. Exemplos de funcionalidades parametrizáveis são: conjuntos de dados e/ou geradores; métodos de aprendizado e/ou detecção de mudanças de conceito; e metodologia de avaliação a ser usada nos experimentos. Em seguida, vem a fase de execução (processamento dos métodos), onde os experimentos são executados. Finalmente, a terceira fase corresponde à análise, onde vários recursos são disponibilizados. Um exemplo é a extração de métricas de avaliação, que inclui o cálculo de testes estatísticos.

Cada um desses recursos, resumidos na Figura 25, está listado abaixo.

Figura 25 – MOAManager Workflow

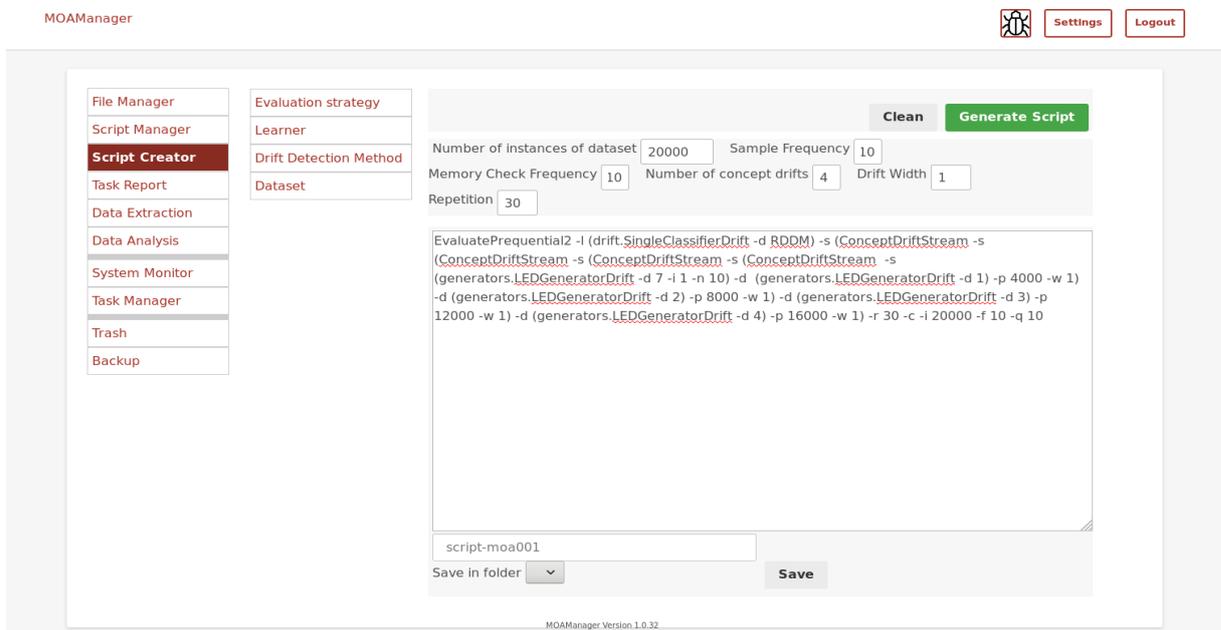


Fonte: O AUTOR (2020).

- **Upload de MOA *Java*TM *Archive* (JAR):** permite que o usuário inclua sua própria versão da área MOA em MOAManager usando arquivos JAR previamente compilados. No entanto, MOAManager já contém uma versão base de MOA em sua distribuição;
- **Editando arquivos Java:** a ferramenta permite editar, adicionar e remover arquivos de código-fonte de um arquivo MOA existente usando uma interface simples com realce de sintaxe;
- **Compilação automática:** depois de alterar os arquivos de código fonte (item anterior), a MOAManager permite gerar automaticamente um novo JAR, possibilitando a execução de novos experimentos usando a versão mais recente do arquivo JAR. Os possíveis erros e avisos de compilação são exibidos ao final da execução;

- **Geração de MOA scripts:** Em uma interface direta e amigável, os experimentos podem ser configurados para serem executados no MOA. Para isso, o usuário deve selecionar a metodologia de avaliação, o classificador, o método de detecção de mudança de conceito e o conjunto de dados e/ou gerador. Neste último, por exemplo, é possível parametrizar o tamanho (em instâncias), o número de mudanças de conceito e a velocidade da mudança de conceito. Em caso de conjuntos de dados do mundo real, o local de cada arquivo deve ser especificado para ser incluído no *script* MOA. Por padrão há alguns exemplos de uso. A Figura 26 apresenta a captura de tela desta funcionalidade da MOAManager;

Figura 26 – Ilustração de como a MOAManager pode ser usada para gerar *Scripts* MOA.



Fonte: O AUTOR (2020).

- **Repetições para calcular intervalos de confiança:** além da geração de *scripts* MOA (item anterior), a MOAManager oferece a opção de repeti-los n vezes, variando as sementes dos geradores de dados, calculando automaticamente os intervalos de confiança de por exemplo, acurácia, uso de memória e tempo de execução de cada *script*. É importante observar que essa opção está disponível apenas para dados artificiais;
- **Criando scripts de execução:** diferentes *scripts* MOA podem ser inseridos sequencialmente em uma *área de texto* para serem executados na ordem em que aparecem. Além disso, também permite que os *scripts* MOA sejam carregados ou salvos para uso futuro e as saídas dos experimentos sejam armazenadas em diretórios previamente definidas pelo usuário. No MOA, por interface de linha comando, apenas um *script* é executado de cada vez, ou seja, não é possível

executá-los em lote. Já na interface gráfica é possível, mas inviável para gerenciar múltiplos experimentos;

- **Definição de processos:** para otimizar o tempo de execução, a MOAManager permite que as experiências sejam executadas usando n processos simultâneos, onde n é um valor definido pelo usuário. Também é possível modificar a quantidade máxima de memória usada por processo Java, permitindo um melhor controle sobre os recursos da máquina;
- **Notificação:** antes de iniciar a execução de um lote de *scripts* MOA, a MOAManager permite que o usuário especifique um endereço de e-mail para receber notificação ao final da execução. Vale ressaltar que, um lote de *scripts* MOA pode ser gerado com base em apenas um arquivo de *scripts* MOA ou múltiplos arquivos;
- **Disponibilidade da máquina:** muitas vezes, os experimentos são executadas em servidores compartilhados. Por esse motivo, a ferramenta MOAManager também fornece informações relacionadas à utilização da memória, processador e do disco rígido para ajudar os usuários a escolherem a configuração mais apropriada para seus experimentos. Um exemplo de configuração a ser definida é o número de processos simultâneos;
- **Contas do usuário:** o acesso aos recursos da MOAManager é restrito e feito sob credencial. Cada usuário possui sua própria área de trabalho e conta com diretórios personalizados para *scripts* MOA, arquivos de resultados, backups e lixeira. Além disso, o administrador do sistema pode definir diferentes privilégios para usuários, limitando alguns recursos críticos da ferramenta. Exemplos de recursos críticos são: compilação; edição de arquivos do código fonte; atualização da ferramenta e/ou recursos, entre outros;
- **Múltiplas Versões do MOA:** MOAManager não restringe seu uso a uma versão particular do MOA. É possível que cada usuário tenha uma ou mais versões associadas a sua conta, incluindo suas próprias modificações e novos códigos.

Na fase de execução, a MOAManager também fornece alguns recursos úteis para auxiliar o usuário, dentre os quais é possível listar:

- **Pausa de experimentos:** a qualquer momento, o usuário pode optar por interromper os experimentos. Essa interrupção inclui o encerramento de todos os processos envolvidos. Além disso, a ferramenta também permite encerrar processos específicos e, dependendo do nível de privilégio da conta, também interromper tarefas de outros usuários (gerenciador de tarefas);

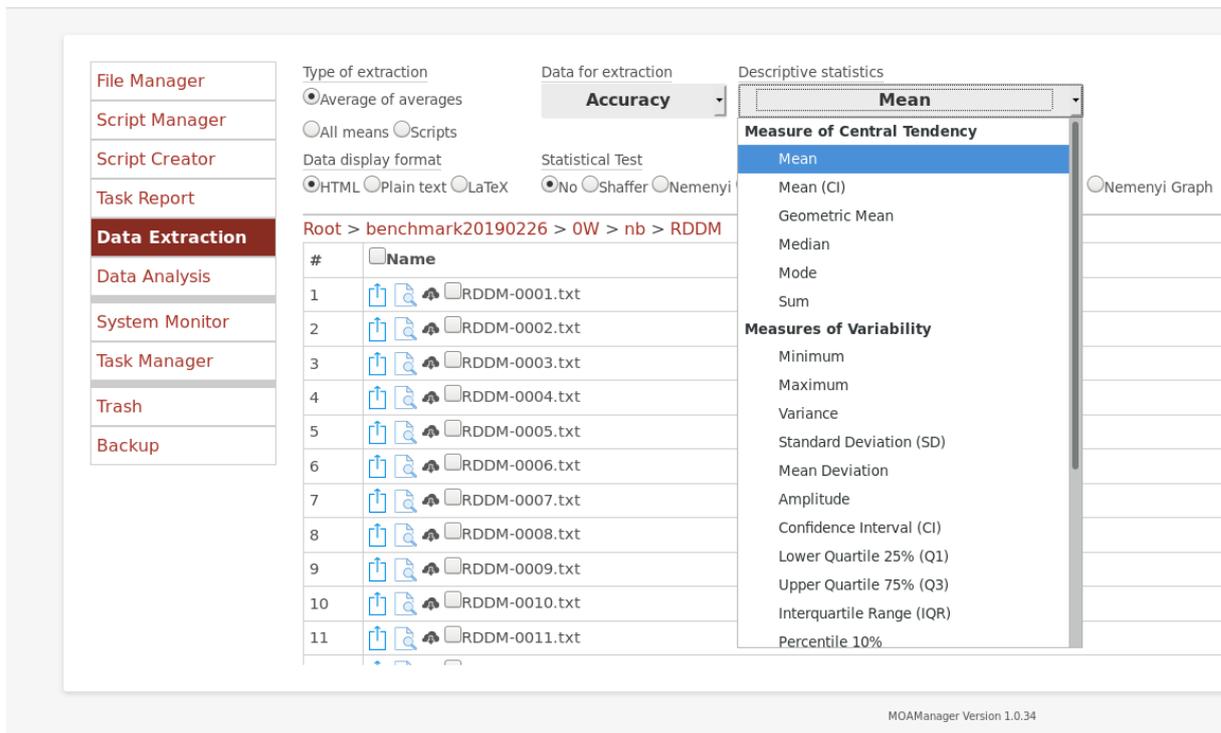
- **Relatório da tarefa:** apresenta uma visão geral da execução dos *scripts* MOA em tempo real. É aqui onde existe a opção de retomada de experimentos;
- **Retomada de experimentos:** em caso de interrupções, voluntárias ou não, a execução de um conjunto de *scripts* MOA pode ser retomada a partir do ponto em que foi interrompida. Geralmente, quanto maior o tamanho do experimento (número de *scripts* MOA) maior o tempo de processamento. Com isso, a probabilidade de utilizar esse recurso é maior, sendo útil para evitar a necessidade de execução dos *scripts* do ponto inicial. Além disso, um outro recurso disponível é o de auxiliar a detecção de possíveis problemas na execução dos *scripts* MOA, sendo possível remover um lote de arquivos inapropriados.

Finalmente, a ferramenta MOAManager também fornece recursos muito úteis para suportar a fase de análise:

- **Extração de resultados:** após a execução de um lote de *scripts* MOA, como descrito anteriormente, o MOA terá criado um grande número de arquivos de saída, um para cada *script*. Abrir esses arquivos, um por um, geralmente consome muito tempo. A ferramenta MOAManager pode rapidamente extrair informações específicas de todos esses arquivos de saída de uma só vez (por exemplo, acurácia média, tempo de execução, etc.) e salvá-los em um único local. A Figura 27 mostra como esta funcionalidade pode ser manipulada;
- **Formatando resultados:** depois da seleção dos resultados desejados, eles são extraídos. Além disso, é possível que o usuário opte por salvá-los em arquivos ou visualizá-los nos seguintes formatos: L^AT_EX, CSV e HTML;
- **Cálculo de métricas para detectores de mudança de conceito:** a partir da extração nos arquivos de saída, a ferramenta MOAManager também obtém cálculo de várias métricas, a exemplo de: componentes da matriz de confusão - FP, FN, TP e TN - acurácia; *recall* (PERRY; KENT; BERRY, 1955); F1 score; MCC (MATTHEWS, 1975); distância média entre as posições reais de mudança de conceito e as posições detectadas nos conjuntos de dados; *Missed Detection Rate* (traduzindo para o português, Taxa de Detecção Perdida) (MDR); *Mean Time to Detection* (traduzindo para o português, Tempo Médio para Detecção) (MTD); *Mean Time between False Alarms* (traduzindo para o português, Tempo Médio Entre Falsos Alarmes) (MTFA); MTR; e detecção de similaridade;
- **Metadados:** além das métricas acima mencionadas, a ferramenta também fornece suporte aos metadados associados aos arquivos de resultados, tais como: função *hash* criptográfica; versão do software; sistema operacional; tempo de processamento utilizado; quantidade de memória principal disponível, etc.

Figura 27 – Ilustra como os dados podem ser extraídos.

MOAManager



Fonte: O AUTOR (2020).

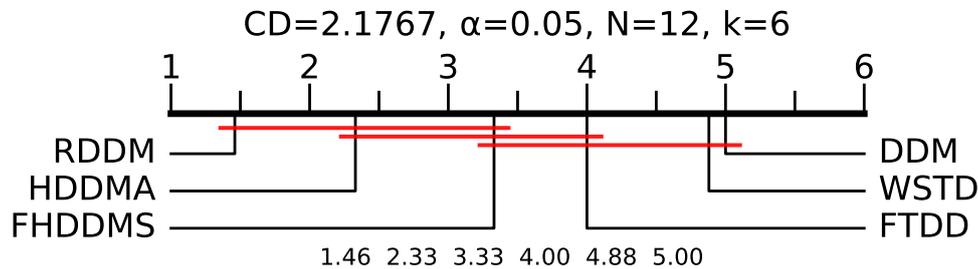
- **Cálculos estatísticos:** outro recurso disponível são os cálculos estatísticos para execuções com mais de uma repetição. O usuário pode optar por incluir seus respectivos intervalos de confiança nos resultados com $p\%$ de confiança. Além disso, se houver a necessidade de realizar comparações entre resultados diferentes (e.g. a precisão de vários métodos) o usuário pode optar por usar o teste de *Friedman* (DEMŠAR, 2006). Neste último, a ferramenta também permite o cálculo dos testes *post-hoc* usando os métodos de *Bonferroni-Dunn*, *Nemenyi*, *Holm*, *Shaffer* e *Bergmann-Hommel* (GARCIA; HERRERA, 2008). Além disso, é possível calcular a distância entre duas métricas usando os métodos *Kullback-Leibler* e *Minkowski*;
- **Criando gráficos:** A partir das estatísticas geradas anteriormente descritas, também é possível criar automaticamente gráficos das diferenças críticas entre os métodos, conforme recomendado por Demšar (2006). A Figura 28 mostra dois exemplos de diagramas de Distâncias Críticas gerados usando MOAManager.

Outros exemplos de recursos suportados por MOAManager são:

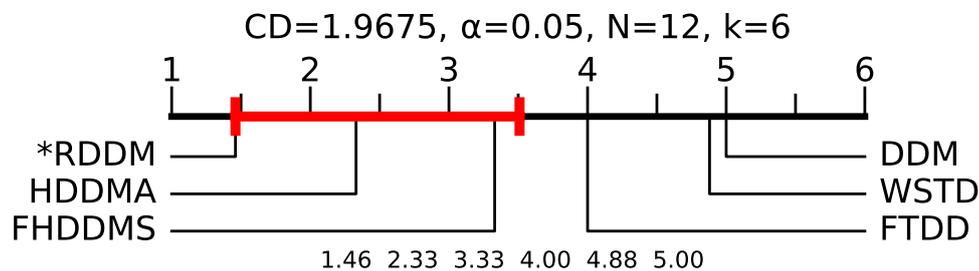
- **Backup:** MOAManager oferece opção de criar arquivo de backup (compressão de dados) contendo arquivos específicos ou todos os arquivos em sua área de trabalho;

Figura 28 – Exemplos de diagramas gerados por MOAManager referentes a resultados de testes estatísticos.

(a) Representação gráfica dos resultados do teste de Friedman e do post-hoc Nemenyi



(b) Representação gráfica dos resultados do teste de Friedman e do post-hoc de Bonferroni-Dunn



Fonte: O AUTOR (2020).

- **Atualização de ferramentas:** para cada nova atualização de versão de ferramenta disponível no repositório *GitHub* mencionado acima, o usuário tem a opção de sincronizar automaticamente sua ferramenta para a versão mais recente. Duas formas de atualização são suportadas: 1) via linha de comando em um terminal; ou 2) via interface gráfica em um navegador. Em ambos os casos, a escolha dos pacotes que serão atualizados fica a critério do usuário;
- **Lixeira:** o primeiro local para onde os arquivos são movidos quando são excluídos. Nesse local, esses arquivos podem ser posteriormente restaurados ou removidos permanentemente.

6.2 IMPLEMENTAÇÃO

MOAManager fornece um ambiente de software de alto nível para executar a maior parte da funcionalidade MOA, abstraindo vários detalhes técnicos. Vale ressaltar que, sua implementação é organizada de acordo com os princípios de orientação a objetos, de modo a facilitar a reutilização componentes.

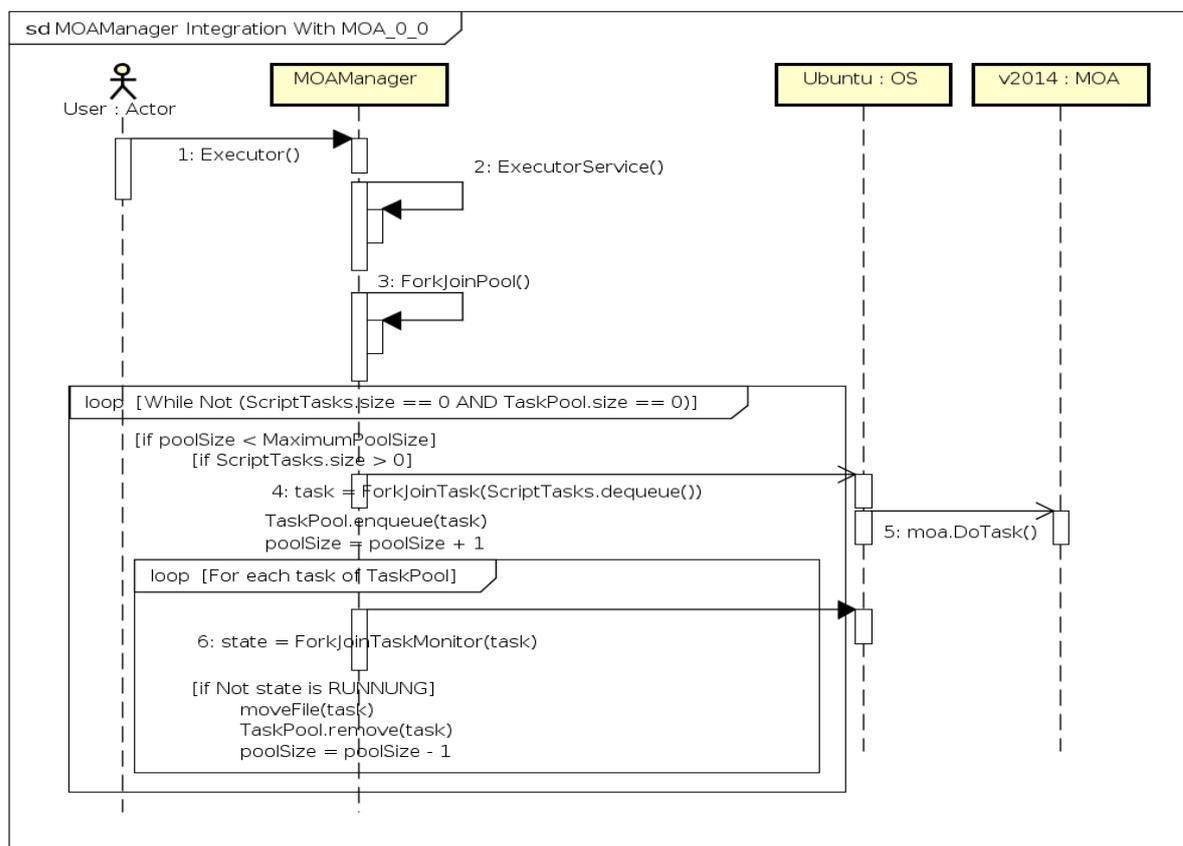
Considerando a natureza interativa do *framework*, foi adotado o padrão arquitetural *Model-View-Controller* (KRASNER; POPE, 1988). Assim, é possível incluir novos componentes, bem como adaptar, modificar ou ampliar a estrutura da ferramenta para

satisfazer as necessidades do desenvolvedor. Além disso, o fato de a MOAManager ser uma ferramenta *web* garante maior portabilidade, com suporte em muitas plataformas.

Seguindo as instruções fornecidas pelo padrão de projeto *Gang of Four* (GAMMA et al., 1994), MOAManager possui padrões diferentes na família de criação. Um deles é o *Singleton*, adotado para controlar as instâncias da classe *application*, localizado dentro do pacote principal. Outro padrão adotado é o *Abstract Factory*, que é usado para encapsular o acesso aos dados do Sistema de Gerenciamento de Banco de Dados (SGBD) em *PDODB.php*. Por fim, em relação ao padrão de família estrutural, o *Facade* foi escolhido já que todo o fluxo de trabalho inicia no arquivo *index.php*, responsável pelas configurações, carregamento de bibliotecas, classes, *framework* e instâncias de aplicação.

A Figura 29 descreve a integração entre MOAManager e MOA. A função **Executor** () fornece os parâmetros necessários para a função **ExecutorService** () iniciar seu trabalho. Essas parametrizações são definidas na funcionalidade **Task Initialize**. A função **ExecutorService** () é responsável por (1) receber e validar os parâmetros; (2) construir a estrutura de diretórios para receber os arquivos de resultados; e (3) criar a lista de execução de tarefas a ser executada por **ForkJoinPool** (), o código que implementa o mecanismo multitarefa.

Figura 29 – MOAManager integração com MOA



Fonte: O AUTOR (2020).

A função **ForkJoinPool** () fornece o ponto de entrada para submissões de trabalhos para **ForkJoinTask** (), bem como gerenciamento de tarefas e operações de monitoramento através de **ForkJoinTaskMonitor** (). A última função verifica o status de execução de uma tarefa, com dois resultados possíveis: **running**, quando a tarefa está em execução; ou **undefined**, quando estiver em um estado indefinido. A função **ForkJoinPool** () também gerencia o problema de simultaneidade, usando o paralelismo, desde que haja disponibilidade de recurso na máquina. Isto permite acelerar as execuções quando existem vários *scripts* MOA que podem ser executados em paralelo.

Note que é recomendado que **ForkJoinPool** () use no máximo um número igual ao de processadores disponíveis para o *Java Virtual Machine* (JVM). O *pool* tentará manter um número de processos ativos igual ao tamanho definido na variável **MaximumPoolSize** variável, adicionando dinamicamente um novo processo sempre que um termina. No entanto, é importante ressaltar que essa funcionalidade não garante o melhor uso dos recursos. O usuário deve levar em conta os recursos e o uso atual da máquina para fornecer uma parametrização apropriada para o **ExecutorService** ().

A integração entre MOAManager e MOA ocorre por meio da função **moa.DoTask** (). Essa é uma interface de linha de comando fornecida pelo MOA para executar *scripts* (sem o uso da *Graphical User Interface* (traduzindo para o português, Interface Gráfica de Usuário) (GUI) padrão) e permite inicializar uma tarefa de execução e aguardar seu resultado (em um arquivo).

Assim, a instrução básica que permite a integração entre MOAManager e MOA contém o *script* MOA de execução e a localização onde o arquivo resultante será armazenado. Finalmente, enquanto o *script* MOA está sendo executado, os resultados parciais são gravados em um arquivo temporário e, quando a execução é concluída, o arquivo completo é movido para a área de armazenamento do usuário.

Embora o gerenciamento de tarefas de execução seja uma importante etapa de integração, a gravação de arquivos de resultados é o principal link de agregação entre as ferramentas. Métricas como precisão de classificação e tempo de execução são coletadas durante a execução do *script* MOA e são usadas para calcular métricas complementares.

Cálculos estatísticos também são usados para definir medidas de centralidade (e.g. média, mediana, etc), variabilidade (e.g. desvio padrão, intervalo de confiança, etc.), simetria e assim por diante.

No MOA, **PrequentialUFPEforDetectors** (), uma versão modificada de **EvaluatePrequential** () que inclui métricas adicionais, é responsável por fornecer os resultados de medidas de desempenho e cálculos estatísticos. Especificamente, as medidas de desempenho fornecidas para classificação são: acurácia, tempo de execução e memória (B/s). Para detecção as métricas disponibilizadas são: FN, FP, TN, TP, MDR, MTD, MTFa, MTR, *precision*, *recall*, MCC e F1 score. Para cada uma destas medidas de desempenho, são calculadas várias informações estatísticas, a exemplo da: \bar{x} média

amostral, intervalo de confiança, média geométrica, mediana, moda, soma, variância, s desvio padrão, desvio médio, mínimo, máximo, amplitude, quartil inferior (Q1), quartil (Q3), intervalo interquartilício (IQR), 10th percentil, 90th percentil e curtose.

Finalmente, a funcionalidade de extração de dados de MOAManager fornece mecanismos para obter métricas considerando múltiplos arquivos e/ou subdiretórios armazenados no mesmo diretório. O conteúdo extraído pode ser visualizado em três formatos: HTML, texto simples e L^AT_EX. Além disso, é importante enfatizar que cada arquivo de resultado gerado por MOAManager também contém (no início) o *script* MOA que executou e gerou os dados. Isso torna fácil para o usuário verificar/identificar os resultados.

6.3 EXPERIÊNCIA DO USUÁRIO

Apesar da popularidade inegável do MOA, a realização de todo o processo experimental usando fluxos de dados em tempo real sempre foi uma tarefa com muitas atividades repetitivas, exigindo várias intervenções manuais e um tempo considerável. Devido a esses fatores, a realização de experimentações massivas (com um grande número de métodos e conjuntos de dados) como os relatados em (BARROS; SANTOS, 2018; BARROS; SANTOS, 2019) seria um processo crítico e não confiável, já que as chances de inserir erros humanos seriam altas.

Por esses e outros motivos surgiu a necessidade de uma ferramenta para dar suporte à automatização de várias dessas atividades manuais, que são de extrema importância. O desenvolvimento da MOAManager foi iniciado em 2015, levando em consideração a ausência de ferramentas com esse propósito.

Assim, o primeiro passo para uma ferramenta que ajudou a reduzir esses problemas foi simplesmente tarefas de scripts MOA para automatizar algumas tarefas específicas. Uma dessas tarefas está relacionada ao cálculo do desempenho dos métodos, com intervalo de confiança e considerando repetições em geradores de dados. Para atingir esse objetivo, diferentes amostras desses geradores de dados devem ser coletadas e, a partir delas, medidas de desempenho são calculadas, o que pode estar associado, por exemplo, à acurácia do método, consumo de memória ou tempo de execução. Assim, o objetivo deste roteiro era coletar automaticamente amostras significativas e realizar o cálculo da média junto com o intervalo de confiança correspondente.

O gerenciamento das execuções de vários métodos em diferentes conjuntos de dados foi outro exemplo de um recurso específico criado. A ideia original era simplesmente passar uma lista de tarefas (*scripts* MOA) para serem executadas de modo que um gerenciador de tarefas a monitorasse e, sempre que uma tarefa terminasse, outra poderia ser iniciada imediatamente. Em outros momentos importantes, este *script* MOA foi atualizado para gerenciar essa fila levando em conta (a) vários arquivos contendo *scripts*

MOA organizados em diretórios e (b) o uso de um número parametrizado de processos na execução, aumentando ainda mais a sua eficiência.

Como membros do grupo de pesquisa começaram a usar as tarefas de automação regularmente, novas demandas surgiram e mais recursos foram criados. Como consequência do crescente número de recursos, nos primeiros meses de 2016 a ferramenta MOAManager foi criada para centralizar esses recursos em uma única ferramenta. Seu objetivo inicial era agregar todos os utilitários existentes de uma forma que pudesse ser acessada por diferentes membros do grupo de pesquisa de forma simples e direta. Esta é a principal razão pela qual uma linguagem de programação *web*, mais especificamente *PHP: Hypertext Preprocessor* (PHP), foi adotada, permitindo que a ferramenta seja acessada de qualquer local e plataforma. A primeira versão da estrutura (arcabouço do sistema) e aplicação integrada da ferramenta estava em funcionamento em abril de 2016.

Após sua primeira versão, a MOAManager continuou a evoluir em várias direções. Recursos relacionados à manipulação de arquivos e diretórios de *scripts* MOA, extração automática de dados em diferentes formatos, pausa e retomada de experimentos, e cálculos estatísticos automáticos foram implementados posteriormente. Em 2018, depois de ser uma realidade dentro do grupo de pesquisa, a MOAManager foi lançado no *github* para torná-la disponível para um grupo maior de usuários, além de permitir possíveis contribuições.

6.4 CONSIDERAÇÕES FINAIS

Esta seção descreveu a MOAManager, uma ferramenta de código aberto destinada a gerenciar o estágio experimental de um projeto em ambientes *on-line*. A ferramenta proposta pode ser usada para suportar diferentes atividades, incluindo a preparação, execução e análise de experimentos usando fluxos de dados. Além de aumentar a produtividade, os recursos disponíveis proporcionam maior confiabilidade para os projetos, uma vez que várias atividades que normalmente seriam realizadas manualmente são totalmente automatizadas.

Uma extensão que poderia ser incorporada em versões futuras da MOAManager é o suporte para fornecer estratégias relacionadas à análise automatizada de dados e execução descentralizada de processos. A ideia principal é permitir que cada cliente processe *scripts* MOA em máquinas diferentes e, após suas execuções, os resultados sejam sincronizados/mesclados em um único local. Outra possibilidade é a criação de *benchmarks* para a avaliação automatizada de novos detectores com base na inclusão de suas implementações de código *Java*. Esses *benchmarks* podem ser disponibilizados publicamente para novos desenvolvedores para orientar seus estudos, contribuindo para a produção de avanços na área de mudança de conceito.

7 CONSIDERAÇÕES FINAIS

Esta seção apresenta as conclusões desta tese a partir das discussões dos resultados delineados nas seções anteriores. São apresentadas as principais contribuições teóricas, bem como as limitações e as sugestões de novas pesquisas a respeito dos temas abordados. Por fim, as contribuições propostas desta tese ao estado da arte da área de mineração de dados em fluxo contínuo de dados, considerando o problema da MC.

Ao longo do tempo para elaboração desta tese, foram realizados diversos experimentos utilizando o MOA. Ao obter novas experiências, novos desafios foram descobertos. Inicialmente, as lacunas no MOA foram a obtenção de métricas específicas para detecção de MC e repetições automáticas das observações. Além dessas, havia também dificuldade em automatizar o processo de execução e extração dos resultados obtidos nos experimentos no MOA. Para suprir essas lacunas, foram implementadas soluções dentro e fora do MOA.

Para gerenciar e minimizar falhas na execução de experimentos em larga escala, foi construída a ferramenta MOAManager. O MOA, apesar de ser uma ferramenta de referência em mineração de fluxos de dados e mudanças de conceito, possui lacunas importantes, como mencionadas anteriormente. A ferramenta MOAManager foi construída com o objetivo de preencher algumas dessas lacunas e potencializar o uso do MOA, principalmente para organizar, executar, extrair e analisar experimentos com muitos dados.

Especificamente em relação aos métodos de DMC, foram propostos e apresentados três métodos na seção 4 e dois na seção 5. Os métodos de comitês DDE, MSDDE, MDDE exploram os métodos convencionais de detecção para otimizar os resultados dos seus membros. A ideia básica é melhorar a eficiência dos resultados levando em consideração diferentes análises de dados. Além dos comitês, os métodos convencionais USDD e USDDX, apresentados na seção 5, também apresentam novas técnicas para DMC.

O DDE (MACIEL; SANTOS; BARROS, 2015) foi inspirado nos comitês de classificadores. A principal motivação do DDE é melhorar os resultados individuais dos seus membros. Uma outra abordagem semelhante, conhecida como *e-Detector* (DU et al., 2014), publicada na mesma época do DDE, obteve resultados experimentais modestos, motivando ainda mais a busca por melhores resultados por meio do DDE. Vale ressaltar que, apesar da simplicidade do DDE, ele ofereceu contribuições ao desempenho de acurácia do modelo com o HT e NB. O DDE pode ser otimizado para melhoria da acurácia *Prequential*, redução de falsos alarmes de detecções e etc.

O MSDDE foi inspirado no DDE e foi motivado pela possibilidade de combinar sinais de *warnings* e drifts dos membros do comitê para detectar MC. Apesar de também ser

um método simples, o MSDDE oferece boas contribuições ao desempenho de precisão do modelo com o HT e NB. O ponto forte do MSDDE é a melhoria em relação a falsos alarmes nas detecções. Além disso, evidências experimentais mostram que o seu desempenho foi o melhor entre os métodos de comitês avaliados.

O MDDE, inspirado também no DDE, foi motivado pela possibilidade de combinar sinais de drifts similares usando entropia da informação e erro médio em janela deslizante. A capacidade de combinar múltiplos sinais de detecções e assumir uma MC utilizando combinações de entropia da informação torna-o robusto para ser utilizado com vários detectores em um mesmo comitê. Assim sendo, essa capacidade de combinar semelhança entre detecções mostrou-se útil para reduzir falsos alarmes.

Os métodos USDD e USDDX foram inspirados no DDM. A principal motivação da técnica do USDD para melhorar os resultados é pressupor convergência do aprendizado para assumir estabilidade do conceito aprendido. A redução elevada na eficiência do classificador até atingir um limiar pré-definido é interpretado como MC. Essa eficiência é calculada usando o protocolo de janela deslizante com dados recentes de tamanho r . A quantidade de dados nessa janela r é parametrizada seguindo indicações do teorema central do limite.

Tratando-se das questões de pesquisa, elas foram respondidas nas seções 4 e 5. Primeiramente, foi possível observar uma melhoria na eficiência dos comitês em relação aos comitês dos trabalhos relacionados. Além disso, também foi possível reduzir falsos alarmes de detecções e maximizar acurácia da classificação *Prequential*. Tratando-se dos métodos de única análise, o USDD e o USDDX também foram capazes de maximizar os resultados em relação a DMC e reduzir falsos alarmes.

Os resultados desses experimentos em larga escala fornecem indicações explícitas dos melhores detectores de MC em termos de: (i) acurácia *Prequential* e detecções; (ii) melhores algoritmos de comitês de detectores, adotando três comitês (C_1 , C_2 e C_3) como base para todos os métodos; e (iii) melhores combinações de detectores nos métodos de comitês. Os experimentos também forneceram a base que tornou possível analisar a eficiência dos detectores, comparando todas as combinações de pares possíveis. Isso contribui para reduzir viés de dados em grupos de resultados que acabam influenciando métricas como média.

Vale ressaltar que os métodos USDD e USDDX estiveram consistentemente entre os melhores resultados de DMC e em termos de acurácia *Prequential*. Além disso, eles foram os métodos que ofereceram as melhores médias de desempenho considerando as diferenças estatísticas para os critérios adotados. Levando em consideração a acurácia *Prequential* em termos individuais, o USDD obteve os melhores resultados com o HT, enquanto o USDDX foi melhor com o NB.

É importante enfatizar que as respostas às questões de pesquisa abordadas nesta tese foram respondidas também nas seções 4 e 5. Em particular, os melhores resultados de

acurácia *Prequential* obtidos por métodos de comitês foram alcançados pelo comitê C_2 , formado pelos detectores FHDDM, FSDD e HDDMA. Além disso, ainda sobre acurácia *Prequential*, o método DDE conseguiu obter melhores resultados em relação aos seus membros. Vale ressaltar, entretanto, que os comitês selecionados para compor a avaliação de desempenho com os métodos do estado da arte não foram selecionados com base apenas na métrica de acurácia *Prequential*.

É importante destacar que, para a avaliação das abordagens propostas, os métodos foram implementados utilizando a linguagem de programação Java dentro do *Framework* MOA. Os experimentos foram realizados considerando diversos cenários e métodos, os quais abrangem: (i) duas velocidades de transições de mudanças de conceito; (ii) seis geradores de dados; (iii) três tamanhos de bases de dados; (iv) sete bases de dados reais; (v) vinte métodos da literatura de mudança de conceito; e (vi) cinco métodos propostos, sendo três de comitês e dois convencionais. Como resultado desta avaliação, verificou-se que, na maioria dos cenários, as abordagens propostas obtiveram resultados competitivos em relação ao estado da arte em mudanças de conceito, apresentando-se como alternativas competitivas.

Assim, conclui-se que o presente trabalho conseguiu atingir seus objetivos, fornecendo soluções para melhoria da acurácia *Prequential* e otimização de falsos alarmes de detecção. Vale ressaltar que as propostas visam preencher uma importante lacuna em função da carência de trabalhos relacionados. Apesar de haver vários métodos para DMC, ainda há um avanço tímido no campo de MC em comparação com aprendizado de máquina.

Até a presente data, foram geradas as seguintes publicações a partir dos resultados obtidos durante a pesquisa e elaboração desta tese:

- Bruno Iran Ferreira Maciel, Silas Garrido Teixeira Carvalho Santos e Roberto Souto Maior Barros. A Lightweight Concept Drift Detection Ensemble. In Proceedings of the 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI) (ICTAI '15). IEEE Computer Society, Washington, DC, USA, 1061-1068. 2015. Publicado (MACIEL; SANTOS; BARROS, 2015);
- Juan Isidro González Hidalgo, Bruno Iran Ferreira Maciel e Roberto Souto Maior Barros. Experimenting with prequential variations for data stream learning evaluation. Experimenting with prequential variations for data stream learning evaluation. Computational Intelligence 35(4): 670-692. 2019. Publicado (HIDALGO; MACIEL; BARROS, 2019);
- Bruno Iran Ferreira Maciel, Silas Garrido Teixeira Carvalho Santos e Roberto Souto Maior Barros. MOAManager: a tool to support data stream experiments. Software: Practice and Experience. Publicado (MACIEL; SANTOS; BARROS, 2020).

7.1 TRABALHOS FUTUROS

Fundamentado na pesquisa desenvolvida nesse trabalho, algumas observações relevantes abrem caminhos para pesquisas futuras, algumas das quais listadas abaixo. Em resumo, previram-se direções para a pesquisa sobre como lidar com as deficiências presentes nas versões dos algoritmos propostos.

- Extensão do DDE e MSDDE: o desempenho da DMC pode ser prejudicado pela escolha dos membros do comitê com deficiências graves pode comprometer o equilíbrio da eficiência do método. Esse desequilíbrio também pode ser aprimorado por escolhas de detectores em cenários não bem apropriados ao seu uso e como exemplo de consequência dessas escolhas está o aumento na taxa de falsos alarmes. O desenvolvimento de um algoritmo de DMC baseado no MSDDE, inspirado no DDE, pode ser uma opção de trabalho futuro;
- Extensão do MDDE: inspirado no DDE, o desenvolvimento foi talvez o mais interessante e inovador dos algoritmos de comitês propostos. A ideia de melhorar as detecções considerando a similaridade entre detecções contraria a lógica utilizada em comitês de classificadores, na qual a dissimilaridade contribui para melhorar os resultados. Logo, explorar a questão da similaridade das detecções, combinando com pesos aos detectores, pode ser outra opção de trabalho futuro;
- Extensão do MDDE: considerar similaridade em detecções recorrentes pode ser mais uma opção de trabalho futuro. Além disso, também há possíveis caminhos em relação ao uso do histórico de detecções individuais e combinadas para prever futuras detecções. Isso pode ser feito considerando um viés probabilístico para correlacionar detecções e membros do comitê;
- Seleção de detectores: explorar a estratégia utilizada no MDDE para inferir similaridade. O erro médio em janela deslizante no momento da DMC é utilizado na estratégia como previsão do rastreamento de DMC similares. Investigar a similaridade para formação de comitês mais eficientes pode ser outro trabalho futuro.
- Comitês de classificadores: comitês de detectores mostraram-se competitivos em relação aos métodos de única análise. Assim sendo, uma opção de investigação é analisar o uso de comitês de detectores em comitês de classificadores com detectores embutidos (SANTOS, 2014; BARROS; SANTOS; GONÇALVES JR., 2016);
- Extensão do USDD: além do tamanho da janela ajustado dinamicamente como no USDDX, pode-se investigar ajustes no critério de decisão usando combinações de parâmetros dinamicamente e histórico de *warning* dentro do mesmo conceito. Além disso, inferir dinamicamente a quantidade mínima necessária para tomadas de decisões também deve ajudar a produzir melhores resultados sob qualquer tipo de MC.

De maneira mais abrangente, uma das direções futuras é desenvolver detectores para tipos específicos de problemas de MC. Além disso, também considerar estimativa de parâmetros em tempo real para otimização dos resultados. Em tais problemas como, por exemplo, graduações longas ou curtas, etc. Há também outra possibilidade de pesquisa futura ao investigar a correlação de métricas e estudar os efeitos ao correlacionar acurácia *Prequential* e falsos alarmes de detecções.

Finalmente, encontrar garantias teóricas para MC é um trabalho futuro importante e desafiador. A relação de DMC com problemas do mundo real pode ser importante para essa direção.

REFERÊNCIAS

- AGRAWAL, R.; IMIELINSKI, T.; SWAMI, A. Database mining: a performance perspective. *IEEE Transactions on Knowledge and Data Engineering*, v. 5, n. 6, p. 914–925, Dec 1993. ISSN 1041-4347.
- BACH, S. H.; MALOOF, M. A. Paired learners for concept drift. In: *Proceedings of 8th IEEE International Conference on Data Mining (ICDM'08)*. Pisa, Italy: [s.n.], 2008. p. 23–32.
- BAENA-GARCIA, M.; CAMPO-ÁVILA, J. D.; FIDALGO, R.; BIFET, A.; GAVALDÀ, R.; MORALES-BUENO, R. Early drift detection method. In: *International Workshop on Knowledge Discovery from Data Streams*. [S.l.: s.n.], 2006. p. 77–86.
- BARROS, R. S. M. *Advances in Data Stream Mining with Concept Drift*. 2017. Professorship (Full) Thesis. Centro de Informática, Universidade Federal de Pernambuco, Brazil.
- BARROS, R. S. M.; CABRAL, D. R. de L.; GONÇALVES JR., P. M.; SANTOS, S. G. T. de C. RDDM: Reactive drift detection method. *Expert Systems with Applications*, v. 90, n. C, p. 344–355, dez. 2017. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417417305614>>.
- BARROS, R. S. M.; HIDALGO, J. I. G.; CABRAL, D. R. de L. Wilcoxon rank sum test drift detector. *Neurocomputing*, v. 275, p. 1954 – 1963, 2018. ISSN 0925-2312. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231217317022>>.
- BARROS, R. S. M.; SANTOS, S. G. T. de C. A large-scale comparison of concept drift detectors. *Information Sciences*, v. 451-452, p. 348 – 370, 2018. ISSN 0020-0255. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0020025518302743>>.
- BARROS, R. S. M.; SANTOS, S. G. T. de C. An overview and comprehensive comparison of ensembles for concept drift. *Information Fusion*, v. 52, p. 213 – 244, 2019. ISSN 1566-2535. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1566253518308066>>.
- BARROS, R. S. M.; SANTOS, S. G. T. de C.; GONÇALVES JR., P. M. A boosting-like online learning ensemble. In: *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN)*. Vancouver, Canada: [s.n.], 2016. p. 1871–1878.
- BENAVOLI, A.; CORANI, G.; MANGILI, F. Should we really use post-hoc tests based on mean-ranks? *Journal of Machine Learning Research*, v. 17, n. 5, p. 1–10, 2016. Disponível em: <<http://jmlr.org/papers/v17/benavoli16a.html>>.
- BERNSTEIN, S. *The Theory of Probabilities*. 1946. Gastehizdat Publishing House, Moscow.
- BETZ, F. *Managing Technological Innovation: Competitive Advantage from Change*. 3. ed. [S.l.]: Wiley-Interscience, 2011.

BIFET, A. *Adaptive Learning and Mining for Data Streams and Frequent Patterns*. Tese (Doutorado) — Universitat Politècnica de Catalunya, 2009.

BIFET, A. Classifier concept drift detection and the illusion of progress. In: RUTKOWSKI, L.; KORYTKOWSKI, M.; SCHERER, R.; TADEUSIEWICZ, R.; ZADEH, L. A.; ZURADA, J. M. (Ed.). *Artificial Intelligence and Soft Computing*. Cham: Springer International Publishing, 2017. p. 715–725. ISBN 978-3-319-59060-8.

BIFET, A.; FRANK, E. Sentiment knowledge discovery in twitter streaming data. In: SPRINGER. *International Conference on Discovery Science*. [S.l.], 2010. p. 1–15.

BIFET, A.; GAVALDÀ, R. Learning from time-changing data with adaptive windowing. In: _____. *Proceedings of the 2007 SIAM International Conference on Data Mining*. [s.n.], 2007. p. 443–448. Disponível em: <<https://epubs.siam.org/doi/abs/10.1137/1.9781611972771.42>>.

BIFET, A.; GAVALDÀ, R. Learning from time-changing data with adaptive windowing. In: *Proc. of 7th SIAM International Conference on Data Mining (SDM'07)*. Minneapolis, MN, USA: [s.n.], 2007. p. 443–448.

BIFET, A.; GAVALDÀ, R. Adaptive learning from evolving data streams. In: *Advances in Intelligent Data Analysis VIII*. Springer Berlin Heidelberg, 2009. p. 249–260. Disponível em: <https://doi.org/10.1007/978-3-642-03915-7_22>.

BIFET, A.; GAVALDÀ, R.; HOLMES, G.; PFAHRINGER, B. *Machine Learning for Data Streams with Practical Examples in MOA*. [S.l.]: MIT Press, 2018. <<https://moa.cms.waikato.ac.nz/book/>>.

BIFET, A.; HOLMES, G.; KIRKBY, R.; PFAHRINGER, B. MOA: massive online analysis. *Journal of Machine Learning Research*, v. 11, p. 1601–1604, 2010. Disponível em: <<http://portal.acm.org/citation.cfm?id=1859903>>.

BIFET, A.; HOLMES, G.; PFAHRINGER, B. Leveraging bagging for evolving data streams. In: *Machine Learning and Knowledge Discovery in Databases*. [S.l.]: Springer, 2010, (LNCS, v. 6321). p. 135–150.

BIFET, A.; HOLMES, G.; PFAHRINGER, B.; KIRKBY, R.; GAVALDÀ, R. New ensemble methods for evolving data streams. In: *Proceedings of 15th ACM International Conference on Knowledge Discovery and Data Mining (KDD'09)*. Paris, France: [s.n.], 2009. p. 139–148.

BIFET, A.; READ, J.; PFAHRINGER, B.; HOLMES, G.; ŽLIOBAITĚ, I. Cd-moa: Change detection framework for massive online analysis. In: TUCKER, A.; HÖPPNER, F.; SIEBES, A.; SWIFT, S. (Ed.). *Advances in Intelligent Data Analysis XII*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 92–103. ISBN 978-3-642-41398-8.

BRZEZIŃSKI, D.; STEFANOWSKI, J. Accuracy updated ensemble for data streams with concept drift. In: CORCHADO, E.; KURZYŃSKI, M.; WOZŃIAK, M. (Ed.). *Hybrid Artificial Intelligent Systems*. [S.l.]: Springer, 2011, (LNCS, v. 6679). p. 155–163.

BRZEZIŃSKI, D.; STEFANOWSKI, J. Reacting to different types of concept drift: The accuracy updated ensemble algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, v. 25, n. 1, p. 81–94, 2013.

- CABRAL, D. R. de L.; BARROS., R. S. M. Concept drift detection based on Fisher's Exact test. *Information Sciences*, v. 442-443, p. 220 – 234, 2018. ISSN 0020-0255. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0020025518301403>>.
- CANO, A.; KRAWCZYK, B. Kappa updated ensemble for drifting data stream mining. *Machine Learning*, Springer Science and Business Media LLC, v. 109, n. 1, p. 175–218, out. 2019. Disponível em: <<https://doi.org/10.1007/s10994-019-05840-z>>.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, n. 3, p. 273–297, Sep 1995. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1023/A:1022627411411>>.
- CORTEZ, P.; CERDEIRA, A.; ALMEIDA, F.; MATOS, T.; REIS, J. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, v. 47, n. 4, p. 547–553, 2009.
- DAWID, A. P. Present position and potential developments: Some personal views: Statistical theory: The prequential approach. *Journal of the Royal Statistical Society. Series A (General)*, JSTOR, p. 278–292, 1984.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, JMLR.org, v. 7, p. 1–30, 2006.
- DIETTERICH, T. G. Ensemble methods in machine learning. In: *Proceedings of the First International Workshop on Multiple Classifier Systems*. London, UK, UK: Springer-Verlag, 2000. (MCS '00), p. 1–15. ISBN 3-540-67704-6. Disponível em: <<http://dl.acm.org/citation.cfm?id=648054.743935>>.
- DOMINGOS, P.; HULTEN, G. Mining high-speed data streams. In: *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2000. (KDD '00), p. 71–80. ISBN 1-58113-233-6. Disponível em: <<http://doi.acm.org/10.1145/347090.347107>>.
- DU, L.; SONG, Q.; ZHU, L.; ZHU, X. A Selective Detector Ensemble for Concept Drift Detection. *The Computer Journal*, v. 58, n. 3, p. 457–471, 06 2014. ISSN 0010-4620. Disponível em: <<https://doi.org/10.1093/comjnl/bxu050>>.
- DUDA, R.; HART, P. *Pattern Classification and Scene Analysis*. [S.l.]: Wiley, 1973.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification (2nd Edition)*. 2. ed. [S.l.: s.n.], 2001. Hardcover. ISBN 0471056693.
- ELWELL, R.; POLIKAR, R. Incremental learning of concept drift in nonstationary environments. *IEEE Transactions on Neural Networks*, v. 22, p. 1517 – 1531, 2011. ISSN 12270581. Disponível em: <<https://ieeexplore.ieee.org/document/5975223/>>.
- FOUNDATION, T. A. S. *Apache SpamAssassin Project*. 2006. <<http://spamassassin.apache.org>>.
- FREUND, Y. Boosting a weak learning algorithm by majority. *Information and Computation*, v. 121, n. 2, p. 256–285, 1995. ISSN 0890-5401. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0890540185711364>>.

- FREUND, Y.; SCHAPIRE, R. E. et al. Experiments with a new boosting algorithm. In: *ICML*. [S.l.: s.n.], 1996. v. 96, p. 148–156.
- FRÍAS-BLANCO, I.; CAMPO-ÁVILA, J. del; RAMOS-JIMÉNEZ, G.; MORALES-BUENO, R.; ORTIZ-DÍAZ, A.; CABALLERO-MOTA, Y. Online and non-parametric drift detection methods based on hoeffding's bounds. *IEEE Transactions on Knowledge and Data Engineering*, v. 27, n. 3, p. 810–823, March 2015. ISSN 1041-4347.
- FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, v. 32, n. 200, p. 675–701, 1937.
- GAMA, J. *Knowledge Discovery from Data Streams*. 1st. ed. [S.l.]: Chapman & Hall/CRC, 2010. ISBN 1439826110, 9781439826119.
- GAMA, J.; CASTILLO, G. Adaptive bayes. In: GARIJO, F. J.; RIQUELME, J. C.; TORO, M. (Ed.). *Advances in Artificial Intelligence — IBERAMIA 2002*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. p. 765–774. ISBN 978-3-540-36131-2.
- GAMA, J.; ŽLIOBAITĚ, I.; BIFET, A.; PECHENIZKIY, M.; BOUCHACHIA, A. A survey on concept drift adaptation. *ACM Computing Surveys*, v. 46, n. 4, p. 44:1–37, 2014.
- GAMA, J.; MEDAS, P.; CASTILLO, G.; RODRIGUES, P. Learning with drift detection. In: *Advances in Artificial Intelligence: SBIA 2004*. [S.l.]: Springer, 2004, (LNCS, v. 3171). p. 286–295.
- GAMA, J.; SEBASTIÃO, R.; RODRIGUES, P. On evaluating stream learning algorithms. *Machine Learning*, Springer, v. 90, n. 3, p. 317–346, 2013.
- GAMMA, E.; HELM, R.; JOHNSON, R.; VLISSIDES, J. *Design Patterns: Elements of Reusable Object-Oriented Software*. [S.l.]: Addison-Wesley, 1994.
- GARCIA, S.; HERRERA, F. An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research*, v. 9, p. 2677–2694, 2008.
- GOMES, H. M.; BIFET, A.; READ, J.; BARDDAL, J. P.; ENEMBRECK, F.; PFAHRINGER, B.; HOLMES, G.; ABDESSALEM, T. Correction to: Adaptive random forests for evolving data stream classification. *Machine Learning*, Springer Science and Business Media LLC, v. 108, n. 10, p. 1877–1878, mar. 2019. Disponível em: <<https://doi.org/10.1007/s10994-019-05793-3>>.
- GONÇALVES JR., P. M.; BARROS, R. S. M. RCD: A recurring concept drift framework. *Pattern Recognition Letters*, Elsevier, v. 34, n. 9, p. 1018–1025, 2013.
- GONÇALVES JR., P. M.; BARROS, R. S. M. Speeding up statistical tests to detect recurring concept drifts. In: LEE, R. (Ed.). *Computer and Information Science*. [S.l.]: Springer, 2013, (Studies in Computational Intelligence, v. 493). p. 129–142.
- GONÇALVES JR., P. M.; SANTOS, S. G. T. de C.; BARROS, R. S. M.; VIEIRA, D. C. L. A comparative study on concept drift detectors. *Expert Systems with Applications*, Elsevier, v. 41, n. 18, p. 8144–8156, 2014.

- HAMED, T.; DARA, R.; KREMER, S. C. Chapter 6 - intrusion detection in contemporary environments. In: VACCA, J. R. (Ed.). *Computer and Information Security Handbook (Third Edition)*. Third edition. Boston: Morgan Kaufmann, 2017. p. 109 – 130. ISBN 978-0-12-803843-7. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780128038437000065>>.
- HAND, D. J.; YU, K. Idiots bayes - not so stupid after all? *International Statistical Review*, v. 69, p. 385–399, 2001.
- HIDALGO, J. I. G.; MACIEL, B. I. F.; BARROS, R. S. M. Experimenting with prequential variations for data stream learning evaluation. *Computational Intelligence*, v. 35, n. 4, p. 670–692, 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/coin.12208>>.
- HO, Y.; PEPYNE, D. Simple explanation of the no-free-lunch theorem and its implications. *Journal of Optimization Theory and Applications*, v. 115, n. 3, p. 549–570, Dec 2002. ISSN 1573-2878. Disponível em: <<https://doi.org/10.1023/A:1021251113462>>.
- HOEFFDING, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, v. 58, p. 13–30, 1963.
- HUANG, D. T. J.; KOH, Y. S.; DOBBIE, G.; PEARS, R. Detecting volatility shift in data streams. In: *2014 IEEE International Conference on Data Mining*. [S.l.: s.n.], 2014. p. 863–868. ISSN 1550-4786.
- HULTEN, G.; SPENCER, L.; DOMINGOS, P. Mining time-changing data streams. In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, USA: [s.n.], 2001. (KDD '01), p. 97–106. ISBN 158113391X. Disponível em: <<https://doi.org/10.1145/502512.502529>>.
- IKONOMOVSKA, E. *Airline dataset*. 2008. Disponível em: <http://kt.ijs.si/elena_ikonovska/data.html>.
- IKONOMOVSKA, E.; GAMA, J.; DŽEROSKI, S. Learning model trees from evolving data streams. *Data Mining and Knowledge Discovery*, v. 23, n. 1, p. 128–168, 2011.
- JAPKOWICZ, J. S. N. *Big Data Analysis: New Algorithms for a New Society*. 1. ed. [S.l.: s.n.], 2016. v. 16. ISBN 978-3-319-26989-4.
- JOHN, G. H.; LANGLEY, P. Estimating continuous distributions in bayesian classifiers. In: MORGAN KAUFMANN PUBLISHERS INC. *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. [S.l.], 1995. p. 338–345.
- KATAKIS G. TSOUMAKAS, I. V. I. *Machine Learning & Knowledge Discovery Group*. 2008. Disponível em: <http://mlkd.csd.auth.gr/concept/_drift.html>.
- KATAKIS, I.; TSOUMAKAS, G.; VLAHAVAS, I. Tracking recurring contexts using ensemble classifiers: an application to email filtering. *Knowledge and Information Systems*, v. 22, n. 3, p. 371–391, 2010.
- KIM, D. H. *Introduction to Systems Thinking Paperback*. 1. ed. [S.l.]: Pegasus Communications, 1999.

KOLTER, J. Z.; MALOOF, M. A. Dynamic weighted majority: An ensemble method for drifting concepts. *Journal of Machine Learning Research*, JMLR.org, v. 8, p. 2755–2790, 2007.

KOTSIANTIS, S. B. Supervised machine learning: A review of classification techniques. In: *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2007. p. 3–24. ISBN 978-1-58603-780-2. Disponível em: <<http://dl.acm.org/citation.cfm?id=1566770.1566773>>.

KRASNER, G. E.; POPE, S. T. A cookbook for using the model-view controller user interface paradigm in smalltalk-80. *J. Object Oriented Program.*, SIGS Publications, Denville, NJ, USA, v. 1, n. 3, p. 26–49, ago. 1988. ISSN 0896-8438. Disponível em: <<http://dl.acm.org/citation.cfm?id=50757.50759>>.

KRAWCZYK, B.; MINKU, L. L.; GAMA, J. ao; STEFANOWSKI, J.; WOŹNIAK, M. Ensemble learning for data stream analysis: A survey. *Information Fusion*, v. 37, p. 132 – 156, 2017. ISSN 1566-2535. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1566253516302329>>.

LARSON, R.; FARBER, B. *Elementary Statistics: Picturing the World*. Fourth. [S.l.]: Pearson, 2010.

LEWIS, D. D. Naive (bayes) at forty: The independence assumption in information retrieval. In: NÉDELLEC, C.; ROUVEIROL, C. (Ed.). *Machine Learning: ECML-98*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. p. 4–15. ISBN 978-3-540-69781-7.

LOSING, V.; HAMMER, B.; WERSING, H. Knn classifier with self adjusting memory for heterogeneous concept drift. In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*. [S.l.: s.n.], 2016. p. 291–300.

MACIEL, B. I. F.; SANTOS, S. G. T. de C.; BARROS, R. S. M. A lightweight concept drift detection ensemble. In: *Proc. of 27th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'15)*. Vietri sul Mare, Italy: [s.n.], 2015. p. 1061–1068.

MACIEL, B. I. F.; SANTOS, S. G. T. de C.; BARROS, R. S. M. Moamanager: A tool to support data stream experiments. *Software: Practice and Experience*, n/a, n. n/a, p. 1–10, 2020. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2795>>.

MANAPRAGADA, C.; WEBB, G. I.; SALEHI, M. Extremely fast decision tree. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, NY, USA: Association for Computing Machinery, 2018. (KDD '18), p. 1953–1962. ISBN 9781450355520. Disponível em: <<https://doi.org/10.1145/3219819.3220005>>.

MARON, O.; MOORE, A. W. Hoeffding races: Accelerating model selection search for classification and function approximation. *Robotics Institute*, p. 263, 1993.

MARRÓN, D.; READ, J.; BIFET, A.; ABDESSALEM, T.; AYGUADÉ, E.; HERRERO, J. R. Echo state hoeffding tree learning. In: *Proceedings of The 8th Asian Conference on Machine Learning*. [S.l.: s.n.], 2016. p. 382–397.

- MATTHEWS, B. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, v. 405, n. 2, p. 442 – 451, 1975. ISSN 0005-2795. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0005279575901099>>.
- MATUSZYK, P.; KREMPL, G.; SPILIOPOULOU, M. Correcting the usage of the hoeffding inequality in stream mining. In: *Advances in Intelligent Data Analysis XII*. Springer Berlin Heidelberg, 2013. p. 298–309. Disponível em: <https://doi.org/10.1007/978-3-642-41398-8_26>.
- MCDIARMID, C. On the method of bounded differences. *Surveys in combinatorics*, v. 141, n. 1, p. 148–188, 1989.
- MINKU, L.; YAO, X. DDD: A new ensemble approach for dealing with concept drift. *IEEE Transactions on Knowledge and Data Engineering*, v. 24, n. 4, p. 619–633, 2012.
- MINKU, L. L. *Online Ensemble Learning in the Presence of Concept Drift*. Tese (Doutorado) — School of Computer Science, The University of Birmingham, July 2010.
- MINKU, L. L.; WHITE, A. P.; YAO, X. The impact of diversity on online ensemble learning in the presence of concept drift. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 5, p. 730–742, 2010.
- MITCHELL, T. M. *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- NISHIDA, K.; YAMAUCHI, K. Detecting concept drift using statistical testing. In: *Proceedings of 10th International Conference on Discovery Science (DS'07)*. [S.l.]: Springer, 2007. (LNCS, v. 4755), p. 264–269.
- OZA, N. C. *Online Ensemble Learning*. Tese (Doutorado) — Berkeley: University of Californi, 2001.
- OZA, N. C.; RUSSELL, S. Online bagging and boosting. In: *Artif. Intellig. and Stat.* [S.l.]: Morgan Kauf., 2001. p. 105–112.
- PAGE, E. S. Continuous inspection schemes. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 41, n. 1-2, p. 100–115, 1954.
- PEARS, R.; SAKTHITHASAN, S.; KOH, Y. S. Detecting concept change in dynamic data streams. *Machine Learning*, v. 97, n. 3, p. 259–293, 2014. ISSN 1573-0565. Disponível em: <<http://dx.doi.org/10.1007/s10994-013-5433-9>>.
- PERRY, J.; KENT, A.; BERRY, M. Machine literature searching X. machine language; factors underlying its design and development. *American Documentation*, v. 6, n. 4, p. 242–254, 1955.
- PESARANGHADER, A.; VIKTOR, H.; PAQUET, E. Reservoir of diverse adaptive learners and stacking fast hoeffding drift detection methods for evolving data streams. *Machine Learning*, v. 107, n. 11, p. 1711–1743, Nov 2018. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1007/s10994-018-5719-z>>.

- PESARANGHADER, A.; VIKTOR, H. L. Fast hoeffding drift detection method for evolving data streams. In: FRASCONI, P.; LANDWEHR, N.; MANCO, G.; VREEKEN, J. (Ed.). *Machine Learning and Knowledge Discovery in Databases*. Cham: Springer International Publishing, 2016. p. 96–111. ISBN 978-3-319-46227-1.
- PRODANOV, C. C.; FREITAS, E. C. d. *Metodologia do trabalho científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico*. 2. ed. Novo Hamburgo, RS: Editora Feevale, 2013. ISBN 978-85-7717-158-3.
- ROBERTS, S. Control chart tests based on geometric moving averages. *Technometrics*, v. 1, n. 3, p. 239–250, 1959.
- ROSS, G. J.; ADAMS, N. M.; TASOULIS, D. K.; HAND, D. J. Exponentially weighted moving average charts for detecting concept drift. *Pattern Recognition Letters*, v. 33, n. 2, p. 191 – 198, 2012. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865511002704>>.
- RUSSELL, S.; NORVIG, P. Inteligência artificial. In: _____. 3. ed. São Paulo, SP, Brasil: GEN LTC, 2013.
- SAKTHITHASAN, S.; PEARS, R.; KOH, Y. S. One pass concept change detection for data streams. In: PEI, J.; TSENG, V. S.; CAO, L.; MOTODA, H.; XU, G. (Ed.). *Advances in Knowledge Discovery and Data Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 461–472. ISBN 978-3-642-37456-2.
- SANTOS, S. G. T. de C. *Adaptable Diversity-based Online Boosting (ADOB)*. 2014. <<http://sites.google.com/site/moamethods/home/adob>>.
- SANTOS, S. G. T. de C.; BARROS, R. S. M. Online adaboost-based methods for multiclass problems. *Artificial Intelligence Review*, 2019. Disponível em: <<https://doi.org/10.1007/s10462-019-09696-6>>.
- SANTOS, S. G. T. de C.; BARROS, R. S. M.; GONÇALVES JR., P. M. Optimizing the parameters of drift detection methods using a genetic algorithm. In: *Proc. of 27th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'15)*. Vietri sul Mare, Italy: [s.n.], 2015. p. 1077–1084.
- SANTOS, S. G. T. de C.; BARROS, R. S. M.; GONÇALVES JR., P. M. A differential evolution based method for tuning concept drift detectors in data streams. *Information Sciences*, v. 485, p. 376 – 393, 2019. ISSN 0020-0255. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0020025519301422>>.
- SANTOS, S. G. T. de C.; GONÇALVES JR., P. M.; SILVA, G.; BARROS, R. S. M. Speeding up recovery from concept drifts. In: *Machine Learning and Knowledge Discovery in Databases*. [S.l.]: Springer, 2014, (LNCS, v. 8726). p. 179–194.
- SCHLIMMER, J. C.; GRANGER, R. H. Incremental learning from noisy data. *Machine Learning*, Springer, v. 1, n. 3, p. 317–354, 1986.
- SEJDIC, E.; FALK, T. H. *Signal Processing and Machine Learning for Biomedical Big Data*. [S.l.]: CRC Press, 2018.

- SIDHU, P.; BHATIA, M. P. S. A novel online ensemble approach to handle concept drifting data streams: diversified dynamic weighted majority. *International Journal of Machine Learning and Cybernetics*, v. 9, n. 1, p. 37–61, Jan 2018. Disponível em: <<https://doi.org/10.1007/s13042-015-0333-x>>.
- TOLLE, J. E. Performance measurement and evaluation of online information systems. In: *Proceedings of the 1985 ACM Thirteenth Annual Conference on Computer Science*. New York, NY, USA: ACM, 1985. (CSC '85), p. 196–203. ISBN 0-89791-150-4. Disponível em: <<http://doi.acm.org/10.1145/320599.320677>>.
- TROMP, J. *UCI Machine Learning Repository*. 1995. Disponível em: <<http://archive.ics.uci.edu/ml>>.
- VAPNIK, V. N.; CHERVONENKIS, A. Y. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, SIAM, v. 16, n. 2, p. 264–280, 1971.
- VAPNIK, V. N.; CHERVONENKIS, A. Y. On the uniform convergence of relative frequencies of events to their probabilities. In: _____. *Measures of Complexity: Festschrift for Alexey Chervonenkis*. Cham: Springer International Publishing, 2015. p. 11–30. ISBN 978-3-319-21852-6. Disponível em: <https://doi.org/10.1007/978-3-319-21852-6_3>.
- WEBB, G. I.; HYDE, R.; CAO, H.; NGUYEN, H. L.; PETITJEAN, F. Characterizing concept drift. *Data Mining and Knowledge Discovery*, v. 30, n. 4, p. 964–994, 2016. ISSN 1573-756X. Disponível em: <<http://dx.doi.org/10.1007/s10618-015-0448-4>>.
- WHITE, M. A.; BRUTON, G. D. *The Management of Technology and Innovation: A Strategic Approach*. 2. ed. [S.l.]: South-Western, 2011.
- WIDMER, G.; KUBAT, M. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, Springer, v. 23, n. 1, p. 69–101, 1996.
- WOLPERT, D. H. The lack of a priori distinctions between learning algorithms. *Neural Computation*, v. 8, n. 7, p. 1341–1390, Oct 1996. ISSN 0899-7667.
- WOLPERT, D. H.; MACREADY, W. G. *No-Free-Lunch Theorems for Search*. [S.l.], 1995.
- WOLPERT, D. H.; MACREADY, W. G. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, v. 1, n. 1, p. 67–82, April 1997. ISSN 1089-778X.
- WOZNAK, M.; KSIENIEWICZ, P.; CYGANEK, B.; WALKOWIAK, K. Ensembles of heterogeneous concept drift detectors - experimental study. In: *Computer Information Systems and Industrial Management - 15th IFIPTC8 International Conference, CISIM 2016, Vilnius, Lithuania, September 14-16, 2016, Proceedings*. [s.n.], 2016. p. 538–549. Disponível em: <http://dx.doi.org/10.1007/978-3-319-45378-1_48>.
- ZHOU, Z.-H.; WU, J.; TANG, W. Ensembling neural networks: Many could be better than all. *Artificial Intelligence*, v. 137, n. 1, p. 239 – 263, 2002. ISSN 0004-3702. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S000437020200190X>>.

ŽLIOBAITĖ, I. Combining time and space similarity for small size learning under concept drift. In: RAUCH, J.; RAŚ, Z. W.; BERKA, P.; ELOMAA, T. (Ed.). *Foundations of Intelligent Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 412–421. ISBN 978-3-642-04125-9.

APÊNDICE A – AVALIAÇÃO DE DESEMPENHO

Uma avaliação de desempenho exige disciplina e comprometimento com as técnicas científicas. Logo, talvez a parte menos intuitiva ao realizar uma avaliação de desempenho seja saber por onde começar. Nesta seção, é apresentada uma metodologia, ou seja, um conjunto de regras e procedimentos para a realização de uma pesquisa, que é válido para estudos que propõem uso de avaliação de desempenho em MC. Vale ressaltar a importância e necessidade de utilizar o método científico e reforçar a ideia desta seção que é de priorização da avaliação de desempenho usada em detectores de MC como forma de concentrar os esforços em questões importantes na área estudada.

No contexto deste trabalho, a avaliação de desempenho envolve principalmente a quantificação de resultados obtidos por um método. Por exemplo, para a natureza deste trabalho é mais interessante comparar os resultados entre diferentes métodos quantitativamente do que qualitativamente. A parte qualitativa será feita com base nos resultados da análise quantitativa.

O restante deste apêndice A está organizada em quatro principais tópicos. No tópico A.1 são apresentadas as considerações iniciais. Seguidamente, no tópico A.2, são apresentados detalhes sobre avaliação de desempenho. Na sequência, no tópico A.3, são apresentadas medidas de avaliação de desempenho utilizadas neste trabalho. Por fim, no tópico A.4 são apresentadas as considerações finais.

A.1 CONSIDERAÇÕES INICIAIS

A ideia de escrever uma seção sobre o assunto de “Avaliação de Desempenho” é importantíssima, principalmente para quem está aprendendo e aplicando experimentos quantitativos (TOLLE, 1985). É compreensível que em trabalhos com limitações de espaço esse assunto receba pouco destaque. Entretanto, não é o caso deste trabalho, sendo assim, pretende-se explicar e detalhar o processo de avaliação de desempenho e medidas comumente adotadas em trabalhos disponibilizados na literatura das áreas de MC e aprendizado de máquina.

Avaliação de desempenho é um processo adotado para mensurar o desempenho e comportamento de técnicas a partir de critérios comuns. Neste trabalho ela é usada como uma ferramenta para conhecer e medir o desempenho dos métodos empiricamente, por meio do confronto entre o que “espera-se” e o que “foi alcançado” por ele. Essa avaliação é um dos recursos mais valiosos para comprovar a relação entre teoria e prática, já que permite diagnosticar e analisar o comportamento de um método pelo cenário que for necessário.

A partir da avaliação de desempenho é possível descobrir as potencialidades de cada

abordagem, bem como os seus pontos fracos e, assim, explorá-los para atingir as melhorias necessárias. Na prática, de forma bastante didática, pode-se dizer que a avaliação de desempenho acaba sendo um meio para identificar problemas dentro de uma abordagem e também uma maneira de valorizar o potencial das qualidades. Além do mais, toda abordagem precisa obter uma apreciação sobre sua proposta e real desempenho. Somente a partir daí, as abordagens poderão desenvolver suas potencialidades e expor suas dificuldades.

Uma análise detalhada do que a/o proposta/método está produzindo (e de que forma) é fundamental, seja para alcançar as metas e/ou fornecer dados importantes sobre a atuação de cada método, subsidiando decisões importantes para escolha dos métodos. Independentemente do resultado (seja ele bom ou ruim), é esse exame que vai ajudar o desenvolvimento de propostas dentro de uma área. Com base nessa avaliação de desempenho, é possível identificar os pontos de melhoria e exemplos a serem seguidos ou alterados, estimulando que sejam produzidos trabalhos cada vez melhores.

Portanto, é importante destacar que a avaliação de desempenho é muito mais do que organizar tabelas/planilhas de resultados. Em particular, geralmente envolve uma atividade em conjunto com os trabalhos para entender pontos positivos e negativos da atuação de cada um deles e, a partir disso, criar um plano para explorar esses quesitos e melhorar os resultados.

Ter uma política de avaliação de desempenho evidencia a preocupação deste trabalho com os demais trabalhos. O desempenho é um fator fundamental na concepção, aquisição e utilização de métodos computacionais. Em outras palavras, obter o melhor desempenho para uma determinada política é um dos desafios na computação e, para alcançá-lo, é necessário ao menos o conhecimento básico da terminologia utilizada na avaliação de desempenho, seus princípios e suas técnicas.

Pesquisadores e demais profissionais das diferentes áreas, geralmente indicam os requisitos de desempenho (objetivos planejados) dos seus métodos/abordagens. É importante avaliar cada proposta e comparar diferentes alternativas para encontrar aquela que melhor atenda as necessidades da política adotada na avaliação de desempenho. Especificamente em detecção no campo de MC, são numerosas as possibilidades de medidas, devido a isso, não existe uma medida padrão de desempenho para detectores, um ambiente de medição padrão (ferramenta, aplicação) ou uma técnica padrão que seja utilizada para todos os casos.

A avaliação de desempenho possui conceitos, técnicas, critérios de avaliação e metodologias sólidas e bem definidas. O uso correto desses recursos permite minimizar ou até evitar erros nas avaliações. Por fim, no próximo tópico A.2, será possível saber o que é, qual a sua importância, como fazer, tipos, métodos, melhores práticas e como explorar para obter seus benefícios. Para tal, algumas medidas de avaliação serão utilizadas neste trabalho para auxiliar em tomadas de decisões.

A.2 AVALIAÇÃO DE DESEMPENHO

É comum que cada trabalho/projeto implemente a política de avaliação de desempenho da forma que mais se adéque as suas expectativas. No intuito de formalizar uma política de avaliação de desempenho, nesta seção são definidas as medidas que serão adotadas no restante do trabalho. É importante ressaltar que essa padronização no uso das métricas de desempenho a serem utilizadas neste trabalho está alinhada com outros trabalhos na literatura da área de aprendizado de máquina e MC.

A fórmula da competência de um método é baseada em comprometimento com as garantias teóricas, habilidades e ação eficaz. Portanto, um método torna-se competente quando faz bem feito o que deve ser feito, não precisa necessariamente ser perfeito, mas tem que ser bem feito. Consequentemente, o instrumento que permite analisar o quão bem feito foi feito é a avaliação de desempenho.

A avaliação de desempenho será feita por um processo sistemático de análise do desempenho de um indivíduo/método em função de requisitos esperados em relação a tarefa do qual vai desempenhar. Neste trabalho, existem dois principais indivíduos a serem observados, a saber, modelo (classificador) e detector. É importante ressaltar que o principal objetivo das avaliações realizadas neste trabalho estejam direcionadas ao detector. Entretanto, um dos papéis do detector é contribuir para a melhoria da tarefa de classificação, devido a isso, é possível observar em trabalhos publicados na literatura da área de MC, uma grande ênfase ao desempenho do modelo nas avaliações de desempenho, principalmente em relação a medida de acurácia da classificação.

Por outro lado, a métrica de acurácia da classificação adotada em avaliações de modelos não apresenta resultado justo em relação ao desempenho do detector. O único compromisso que será avaliado neste trabalho em relação aos modelos corresponde à medida de acurácia da classificação. Por outro lado, o compromisso dos detectores será maior, pois o foco deste trabalho é sobre detectores.

Serão estabelecidos os objetivos planejados (compromisso) para os detectores. A principal ideia da avaliação é identificar em que pontos ele se afastou ou aproximou-se dos resultados dessas metas que foram estabelecidas. Quanto mais próximo estiver o detector dos seus objetivos melhor será avaliado, isto é, melhor será seu desempenho.

Mas por que avaliar o desempenho dos métodos? É uma pergunta que exige uma resposta objetiva, apesar de não ser simples e direta respondê-la. Para responder, são apresentados pontos de vista, que visam representar bem a ideia da avaliação e para isso são definidos quatro principais pontos, a saber:

- Medir a entrega do método (podendo aqui ser o modelo ou detector) – refere-se a quanto ele conseguiu atingir seus objetivos;
- Subsidiar decisões como:

- Remoção, isto é, permitir a exclusão do método de novos experimentos, devido a seu desempenho insatisfatório ou inferior a outros.
 - Promoção, isto é, possibilita qualificar satisfatoriamente o método de acordo com os objetivos alcançados. Assim como permitir explorar novos objetivos mais complexos ou cenários diferentes;
 - Realocação, isto é, permite ajustar as qualidades do método para situações no qual ele possa obter melhor desempenho. Possibilitando sugerir que dado método possa obter melhores resultados em dada situação/cenário.
- Informar detalhes – na medida que é avaliado o desempenho do método, informações sobre os objetivos planejados são obtidas, permitindo melhorar a compreensão em relação ao comprometimento esperado pelo método;
 - Suprir o processo de ajustes e desenvolvimento – o método é avaliado continuamente e a partir dessas avaliações é possível realizar ajustes que melhoram seus resultados em relação aos objetivos planejados.

Independentemente do ponto de vista, a interpretação dos resultados é a análise propriamente dita que interessa ao avaliador. A seguir são apresentadas algumas técnicas utilizadas em avaliações e desempenho.

A.3 MEDIDAS DE AVALIAÇÃO DE DESEMPENHO

De acordo com White e Bruton (2011, p.40), o desenvolvimento econômico de longo prazo surge do progresso tecnológico, onde novas indústrias são estabelecidas para implementar novos sistemas econômicos. De modo que a busca pelo desempenho seja um dos principais fatores para impulsionar o progresso tecnológico. Essa busca acontece por diferentes pontos de vista, dentre eles o econômico, no qual a intenção é obter o maior desempenho com o menor custo.

Definição 6 – Sistema – qualquer grupo de partes interdependentes que interagem entre si e se inter-relacionam, formando um todo complexo e unificado com um propósito específico (KIM, 1999, p.4). É importante ressaltar que sem essas interdependências, existe apenas uma coleção de partes, não um sistema.

Segundo Betz (2011, p.39), o progresso nos campos tecnológico e desenvolvimento econômico não são suaves e contínuos e ambos envolvem processos de descontinuidade (seja tecnológico) e periodicidade (ciclos econômicos). Isso ocorre também porque o progresso tecnológico é finito em uma tecnologia qualquer e sendo assim, isso causa impactos nas indústrias, porque as estruturas industriais são criadas a partir de novas tecnologias básicas.

De forma geral, uma avaliação de desempenho representa um conjunto de técnicas e regras que visam a sistematização no processo de mensuração do desempenho do sistema. A implementação de uma avaliação de desempenho é feita com base no entendimento do sistema/método a ser analisado e das técnicas de avaliação de desempenho existentes. A análise de desempenho pode ser feita devido a diferentes pontos de vista, dentre os quais é possível destacar os pontos observados por este trabalho, a saber:

- **Comparar alternativas** - a comparação de alternativas consiste em escolher entre as alternativas, a que melhor ajusta-se aos requisitos estabelecidos na política de avaliação de desempenho. Em termos de requisitos podem ser: maior percentual de acurácia da classificação; menor tempo de processamento; menor consumo de memória; menor quantidade de falsos positivos e falsos negativos;
- **Identificar pontos fortes e fracos** - neste caso, métodos normalmente possuem limitações e na medida em que as limitações são identificadas, novas soluções podem ser exploradas para oferecer melhorias e conseqüentemente refletir em um melhor desempenho. Exemplos de limitações podem ser associadas a redução do desempenho em determinadas bases de dados artificiais e/ou modelo de classificação utilizado;
- **Determinar configuração ideal dos parâmetros de um método** - determinar a configuração ideal de um método permite maximizar o desempenho dele. Por exemplo, para cada cenário, seja real ou artificial é possível a partir do ajuste de parâmetros encontrar configurações ideais. Isso pode ser alcançado manualmente ou por maneiras automatizadas;
- **Fazer previsão de desempenho em função do tamanho do conceito** - analisar o desempenho do método em função do crescimento do conceito é usado para inferir o desempenho do método a partir de simulações com diferentes tamanhos de conceitos e analisar a condição de melhoria ou piora no desempenho na medida que os conceitos mudam de tamanho.

Neste sentido, são utilizadas diferentes métricas com cálculos distintos. Basicamente, essas métricas procuram representar quantitativamente o desempenho do método em relação aos números de FN, FP, TN e TP para um dado cenário. Dado um experimento em que pretende-se avaliar o desempenho do detector, espera-se que o detector seja capaz de detectar uma MC corretamente.

O Quadro 3, representa uma matriz de confusão, na qual apresenta os possíveis resultados da análise do detector. Ela mostra as frequências de classificação para cada ocorrência de estimativa de detecção. São apresentados em “Valores Reais” as classificações esperadas e em “Valores Estimados” são mostradas as classificações

Quadro 3 – Matriz de confusão com resultados de detecção

| | | Valores Estimados | |
|---------------|--------------------|---|---|
| | | Drift (Positivo) | Não Drift (Negativo) |
| Valores Reais | Drift (Verdadeiro) | TP (Drift Existe e Estimativa diz que Existe) | FN (Drift Existe e Estimativa diz que Não Existe) |
| | Não Drift (Falso) | FP (Drift Não Existe e Estimativa diz que Existe) | TN (Drift Não Existe e Estimativa diz que Não Existe) |

Fonte: O AUTOR (2020).

estimadas pelo detector. Aproveitando o exemplo do Quadro 3, as métricas de TP, TN, FP e FN representam as frequências para cada situação, a saber:

- TP: a mudança de conceito foi analisada corretamente pelo detector, pois a mudança de conceito existe e foi detectada;
- TN: a mudança de conceito foi analisada corretamente pelo detector, pois a mudança de conceito não existe e não foi detectada;
- FP: a mudança de conceito foi analisada incorretamente pelo detector, pois a mudança de conceito não existe, mas foi detectada;
- FN: a mudança de conceito foi analisada incorretamente pelo detector, pois a mudança de conceito existe e não foi detectada.

As Tabelas 16 e 17 apresentam dois cenários, um em cada. Na primeira, Tabela 16, são apresentados dados de uma situação na qual as detecções estimadas foram todas sem erros, isto é, $TP = 4$, $FP = 0$, $FN = 0$ e $TN = 9996$. Na Tabela 17, os dados mostram uma situação que houveram estimativas erradas de detecção, podendo ser visto em $TP = 3$, $FP = 2$, $FN = 1$ e $TN = 9994$.

Tabela 16 – Base de dados com 10k instâncias e 4 drifts reais - situação ideal

| | | Valores Estimados | |
|---------------|-----------|-------------------|-----------|
| | | Drift | Não Drift |
| Valores Reais | Drift | 4 | 0 |
| | Não Drift | 0 | 9996 |

Fonte: O AUTOR (2020).

Sendo assim, o detector da Tabela 16 obteve o seguinte desempenho:

- TP - Estimou a **mudança de conceito** 4 vezes corretamente - total 4;

- TN - Deixou de estimar a **mudança de conceito** 9996 vezes corretamente - total 9996;
- FP - Estimou a **mudança de conceito** 0 vez incorretamente;
- FN - Deixou de estimar a **mudança de conceito** 0 vez incorretamente.

Tabela 17 – Base de dados com 10k instâncias e 4 drifts reais

| | | Valores Estimados | |
|---------------|-----------|-------------------|-----------|
| | | Drift | Não Drift |
| Valores Reais | Drift | 3 | 1 |
| | Não Drift | 2 | 9994 |

Fonte: O AUTOR (2020).

Sendo assim, o detector da Tabela 17 obteve o seguinte desempenho:

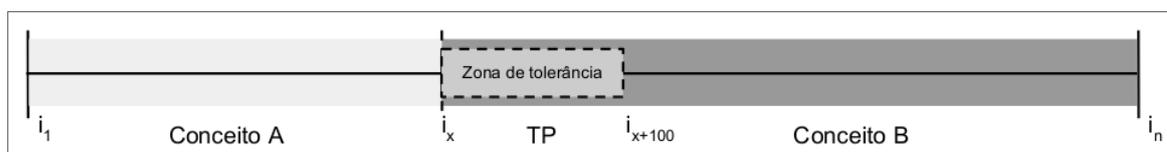
- TP - Estimou a **mudança de conceito** 3 vezes corretamente - total 4;
- TN - Deixou de estimar a **mudança de conceito** 9994 vezes corretamente - total 9996;
- FP - Estimou a **mudança de conceito** 2 vezes incorretamente;
- FN - Deixou de estimar a **mudança de conceito** 1 vez incorretamente.

Em relação aos detectores, existem diferentes estratégias para a construção deles. A abordagem mais tradicional é reativa, onde a análise do fluxo de dados é feita incrementalmente, sob à demanda de dados classificados. Nesse tipo de abordagem, os detectores reagem aos fatos analisados no fluxo de dados, isto é, uma legítima detecção é estimada após uma MC real acontecer.

Na área de detecção é aceitável adotar uma zona de tolerância para considerar uma estimativa de identificação como correta. Para avaliar o desempenho dos detectores, este trabalho adotou a zona de tolerância de 100 instâncias após a mudança de conceito ocorrer, mesmo tamanho de tolerância adotado em outros trabalhos (BARROS; SANTOS, 2018; BARROS; HIDALGO; CABRAL, 2018). A Figura 30 apresenta um exemplo de zona de tolerância entre dois conceitos, A e B, no qual o conceito A encerra em i_x , por sua vez a zona de tolerância tem intervalo de i_x até i_{x+100} e caso alguma estimativa de mudança ocorra dentro desse intervalo de tolerância, a estimativa é considerada como correta.

Na Tabela 18, os dados mostram uma situação no qual houve estimativas erradas de detecção, podendo ser visto em $TP = 2$, $FP = 3$, $FN = 1$ e $TN = 9994$. Sendo assim, o detector da Tabela 18 obteve o seguinte desempenho:

Figura 30 – Zona de tolerância na detecção



Fonte: O AUTOR (2020).

Tabela 18 – Base de dados com 10k instâncias e 4 drifts reais - dados estimados e reais de mudança de conceito. A classificação corresponde a categoria na qual a identificação do drift foi realizada, pertencendo a um conjunto de 3 classes, a saber: TP, FP e FN. A classificação para TN corresponde ao total de instâncias na base de dados menos o total de TP + FP + FN.

| | | | | | | |
|-------------------|-------|-------|-------|-------|-------|-------|
| Valores Estimados | | 4.110 | 6.020 | 8.070 | 9.000 | 9.800 |
| Valores Reais | 2.000 | 4.000 | 6.000 | 8.000 | | |
| Classificação | FN | FP | TP | TP | FP | FP |

Fonte: o autor.

Tabela 19 – Matriz de confusão com dados da Tabela 18

| | | Valores Estimados | |
|---------------|-----------|-------------------|-----------|
| | | Drift | Não Drift |
| Valores Reais | Drift | 2 | 1 |
| | Não Drift | 3 | 9994 |

Fonte: O AUTOR (2020).

- TP - Estimou a **mudança de conceito** 2 vezes corretamente nos pontos [6.020 e 8.070] - o total correto esperado eram de 4;
- TN - Deixou de estimar a **mudança de conceito** 9.994 vezes corretamente - o total correto esperado era de 9.996;
- FP - Estimou a **mudança de conceito** 3 vezes incorretamente nos pontos [4.110, 9.000 e 9.800];
- FN - Deixou de estimar a **mudança de conceito** 1 vez incorretamente no ponto [2.000].

Ainda em relação a Tabela 18 é possível fazer mais algumas considerações, a saber:

- Total de estimativas corretas de mudanças de conceito pode ser expresso por: TP + TN. Em caso de 100% de acertos, será igual ao tamanho da base de dados, por exemplo, [TP=4 + TN=9.996] = 10k. No caso particular dos dados da Tabela 18, fica [TP=2 + TN=9.994] = 9.996 análises corretas;

- Total de estimativas incorretas de mudanças de conceito pode ser expresso por: $FP + FN$. Em caso de 100% de acertos, será igual a zero, por exemplo, $[FP=0 + FN=0] = 0$. No caso particular dos dados da Tabela 18, fica $[FP=3 + FN=1] = 4$ análises incorretas;
- Total de estimativas de mudanças de conceito pode ser expresso por: $TP + FP$. No caso particular dos dados da Tabela 18, fica $[TP=2 + FP=3] = 5$ detecções, sendo 2 corretas e 3 incorretas;
- Total de estimativas na qual não ocorrem mudanças de conceito pode ser expresso por: $TN + FN$. Em caso de 100% de acertos, será igual ao tamanho da base de dados menos a quantidade de mudanças de conceito reais, por exemplo, $[TN=9.996 + FN=0] = 9.996$, isto é, $10k - 4 = 9.996$. No caso particular dos dados da Tabela 18, fica $[TN=9.994 + FN=1] = 9.995$, isto é, foram feitas 9.995 análises na qual o detector estimou não existir mudança de conceito;
- Se o detector tiver um alto valor para FN, ele não estima corretamente as mudanças de conceito, ou seja, sua sensibilidade é baixa. Por outro lado, se FP obtiver um alto valor ele terá sua sensibilidade muito alta, pois vai está estimando mudanças de conceito em momentos no qual não existem ou que estejam fora da zona de tolerância.

Nessa perspectiva sobre acertos e erros de detecção, o desempenho dos detectores pode ser obtida analisando o número de falsos negativos e falsos positivos. Um FN ocorre quando a MC existe, mas não é detectado, na estatística é conhecido como Erro Tipo 2. Por outro lado, um FP ocorre quando a MC é detectada, mas ela não existe (Erro Tipo 1). Para um detector, um falso negativo é pior que um falso positivo porque a MC ocorreu, mas não foi identificada pelo detector.

Para concluir o raciocínio em relações aos FPs e FNs, a Figura 31 ilustra duas situações. Na primeira parte, com legenda de Erro Tipo 1, um homem recebe o diagnóstico que está grávido. Na segunda parte, com legenda de Erro Tipo 2, uma mulher grávida recebe o diagnóstico que não está grávida. A principal ideia da ilustração é explicar os erros de FP (Erro Tipo 1) e FN (Erro Tipo 2), que são componentes essenciais da avaliação das estimativas de detecções.

Um bom detector, em princípio, deve obter uma análise rápida (embora arbitrária) da MC com um pequeno erro de variação entre as posições de deslocamento estimados e reais. Cada projeto de detector é mais adequado para uma situação, contexto, isto é, dependente das características dos dados e como elas vão enviar o indutor. Por exemplo, para um detector que precisa identificar possíveis falhas em um sistema ou em teste para detectar doença, geralmente é essencial detectar a MC o mais rápido possível (pouco atraso na detecção), enquanto o número de detecções incorretas deve ser pequeno.

Figura 31 – Error tipo 1 e erro tipo 2



Fonte: O AUTOR (2020).

O objetivo intrínseco dos detectores é minimizar simultaneamente a taxa de detecções incorretas e o atraso nas detecções. Outras medidas interessantes são a acurácia da estimativa em relação a posição da MC, a variabilidade entre as detecções e etc. Nas próximas subseções são apresentadas algumas das métricas utilizadas em trabalhos na literatura de área de MC que são utilizadas neste trabalho.

A.3.1 Taxa de Sucesso ou Exatidão ou Acurácia (Accuracy)

A proporção de estimativas corretas para o total de elementos estimados pode ser expressa por (HAMED; DARA; KREMER, 2017):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

A.3.2 Taxa de Insucesso ou Erro (Err)

A proporção de estimativas incorretas para o total de elementos estimados pode ser expressa por (HAMED; DARA; KREMER, 2017):

$$\text{Error} = \text{Err} = \frac{FP + FN}{TP + FP + TN + FN}$$

ou pode ser obtido por:

$$\text{Err} = 1 - \text{Accuracy}$$

A.3.3 Precisão (Precision)

A proporção de verdadeiros positivos estimados pode ser expressa por (HAMED; DARA; KREMER, 2017):

$$\text{Precision} = \frac{TP}{TP + FP}$$

A.3.4 MCC

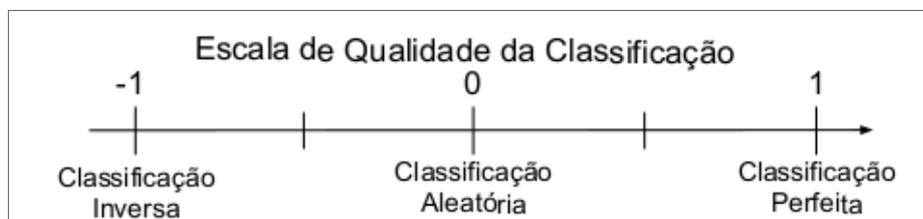
O cálculo do MCC (MATTHEWS, 1975) é usado em aprendizado de máquina como uma medida da qualidade das classificações. O resultado do cálculo do MCC varia entre -1 e 1, no qual 1 (máximo valor positivo) representa uma previsão perfeita, 0 uma previsão aleatória média e -1 (mínimo valor negativo) uma previsão inversa. É definida da seguinte forma:

$$\text{MCC} = \frac{(TP \times TN) - (FN \times FP)}{\sqrt{(TP + FN) \times (TN + FP) \times (TP + FP) \times (TN + FN)}}$$

O cálculo leva em conta os TP, TN, FP e FN e geralmente é considerado como uma medida balanceada que pode ser usada mesmo se as classes forem de tamanhos muito diferentes. O MCC é, em essência, um valor de coeficiente de correlação entre valores observados e previstos.

A Figura 32, apresenta uma escala de qualidade das classificações dos resultado do MCC. Quanto mais próximo de 1, melhor a qualidade da previsão do detector e quanto mais próximo de -1, pior a qualidade, logo um maior desacordo entre previsões e observações.

Figura 32 – Escala do Coeficiente de Correlação de Matthews



Fonte: O AUTOR (2020).

A.3.5 MTFA

O MTFA caracteriza a frequência com que são detectados os FP (falsos alarmes) quando não há alteração. A taxa de FP é definida como $1/\text{MTFA}$. Um bom detector de MC teria um alto MTFA. Vale ressaltar que o “tempo” em problemas tratados neste trabalho de doutorado está relacionado a ordem da posição em que a instância aparece.

A.3.6 MTD

O MTD caracteriza a reatividade do sistema a mudanças depois que elas ocorrem. Um bom detector de mudança teria um pequeno MTD.

A.3.7 MDR

A MDR dá a probabilidade de não receber um alarme, isto é, um FN, em outras palavras quando houver MC e não ser detectada. É calculado com base na fração de alterações não detectadas em todas as alterações que ocorreram. Um bom detector teria MDR pequeno ou igual a zero.

A.3.8 ARL(θ)

O *Average Run Length* (traduzindo para o português, Média de Duração da Execução) (ARL(θ)) generaliza sobre o MTFA e MTD. Ele quantifica quanto tempo tem-se que esperar antes de detectar uma MC de tamanho θ na variável monitorada (posição esperada).

$$ARL(\theta = 0) = MTFA, \quad ARL(\theta \neq 0) = MTD$$

Para fazer uma comparação justa dos detectores de mudanças, a estrutura de avaliação precisa conhecer as mudanças verdadeiras dos dados para avaliação dos algoritmos de detecção de mudanças. Assim, é preciso usar conjuntos de dados sintéticos com a verdade básica. Antes que uma verdadeira mudança aconteça, todos os alarmes são considerados como FP (falsos alarmes). Depois que uma verdadeira mudança (isto é, um TP) ocorre, a primeira detecção que é sinalizada é considerada como o verdadeiro alarme TP. Depois disso e antes que ocorra uma nova mudança verdadeira, as detecções consequentes são consideradas como falsos alarmes. Se nenhuma detecção for sinalizada entre duas mudanças verdadeiras, então é considerada uma detecção perdida.

A.3.9 MTR

A MTR (BIFET et al., 2013) avalia a qualidade no comprometimento justo entre detecção rápida e FP. É definida da seguinte forma:

$$MTR(\theta) = \frac{MTFA}{MTD} \times (1 - MDR) = \frac{ARL(0)}{ARL(\theta)} \times (1 - MDR)$$

MTR representa a razão entre o *tempo médio e alarmes falsos* e o *tempo médio de detecção*, multiplicado pela probabilidade de detectar um alarme. Um algoritmo de detecção de mudança ideal teria uma taxa de falsos positivos baixa (o que significa um tempo médio alto entre falsos alarmes), um tempo médio baixo para detecção e uma baixa taxa de detecção perdida. A comparação de dois detectores de mudança para uma mudança específica θ é fácil com esta medida: o algoritmo que possui o maior valor de MTR(θ) deve ser preferido.

A.4 CONSIDERAÇÕES FINAIS

Pode-se entender pelo termo política, como “orientação ou conjunto de regras a serem seguidas” na avaliação de desempenho. Primeiramente, são utilizadas medidas para avaliação de classificação binária. Caso particular que interessa as detecções de MC, dado pelo conjunto formado por duas categorias, a saber: mudança de conceito ou não mudança de conceito. São adotados três grupos de medidas de desempenho, a saber:

- Critérios isolados - calculado usando apenas uma medição;
- Critérios pareados - calculado usando duas medições;
- Critérios Combinados - calculado usando mais de duas medições.

No grupo dos critérios isolados estão TP, TN, FP, FN, MTFa, MTD e MDR. Em critérios pareados estão Precisão e ARL. Por fim, as medidas de Acurácia, Erro, MCC e MTR ficam agrupadas no grupo de critérios combinados. Assume-se que essas medidas serão calculadas para todos os casos experimentados neste trabalho, todavia, vale ressaltar que essas medidas são para avaliar o detector.

Primeiramente, é necessário entender que os principais objetivos deste trabalho de doutorado são sobre detecção de MC e a necessidade dos resultados esperados são sobre esses objetivos. Neste sentido, segundo Bifet (2017), não é uma boa metodologia de avaliação de detectores usar apenas acurácia da classificação, pois pode existir dependência temporal dos dados e um classificador sem detector pode superar um classificador com detector. Em Bifet et al. (2013), também é discutido sobre isso e nesse trabalho ele propõe uma nova maneira para avaliar os detectores em relação aos resultados esperados por meio de uma nova medida de desempenho que monitora o comprometimento entre detecção rápida e FP.

Intuitivamente, quando um novo detector é proposto, uma maneira comum de mostrar os benefícios dele é usando um classificador e realizada uma avaliação onde cada vez que ocorre uma mudança, o classificador principal é substituído por um novo. Entretanto, é preciso levar em consideração a existência de outras possibilidades para avaliação do desempenho de detectores e cada possibilidade pode oferecer diferentes visões de resultados. Assim sendo, para cada experimento, deseja-se obter no mínimo os cálculos das médias de acurácia *Prequential* do Modelo e MCC das detecções.

Para avaliar o desempenho das medidas serão utilizados testes estatísticos Não Paramétricos para avaliar a nulidade da hipótese. Em outras palavras, o teste estatístico deve avaliar se as amostras pertencem a mesma população, isto é, se elas são estatisticamente equivalentes. Para verificar o desempenho das amostras em termos estatísticos será utilizado como requisito mínimo o teste de *Wilcoxon signed-rank*.

As medições serão realizadas por meio do MOA *Framework* e pacote de medição desenvolvido pelo grupo de alunos sob supervisão do Prof. Dr. Roberto Souto Maior de Barros da UFPE. Os experimento foram gerenciados pela ferramenta MOAManager, responsável pela comunicação com MOA e gerência da execução dos scripts MOA e resultados.

APÊNDICE B – DETECTORES CONVENCIONAIS

Nesta seção são avaliados os métodos de DMC de única análise do estado da arte em MC apresentados na seção 2. São avaliados os desempenhos de dezesseis métodos detectores: DDM, EDDM, FHDDM, FHDDMS, FPDD, FSDD, FTDD, HDDMA, HDDMW, RDDM, SEED, STEPDP, SeqDrift1, SeqDrift2 e WSTD. Os resultados experimentais mostram que o método FSDD obteve os melhores resultados em comparação com os métodos do estado da arte comparados, considerando a média de desempenho nos critérios acurácia da classificação, MCC e MTR da detecção. Tratando-se de um trabalho com tema em MC, adotou-se os dois critérios de avaliação de DMC, ambos robustos para avaliar o desempenho de DMC e utilizando em outros trabalhos como Barros, Hidalgo e Cabral (2018), Cabral e Barros. (2018), Bifet (2017): o MTR avalia o detector em mudanças abruptas e graduais; enquanto o MCC é adotado para avaliar apenas os casos em que ocorrem mudanças de conceito abruptas. O restante desta seção está organizada da seguinte maneira: no tópico B.1 são apresentadas as considerações iniciais e explicações em relação aos objetivos esperados da avaliação; seguidamente, B.2 apresenta a motivação para elaboração da avaliação de desempenho; na sequência a B.3 mostra e discute os resultados; por fim, a seção B.5 apresenta as considerações finais sobre os resultados obtidos.

B.1 CONSIDERAÇÕES INICIAIS

A partir da avaliação de desempenho com os métodos apresentados na seção 2, pretende-se identificar, classificar e ordenar os métodos detectores de acordo com o desempenho individual e utilizá-los para comparar com os métodos propostos por este trabalho. Assim sendo, as abordagens propostas nesta tese serão avaliadas em relação aos melhores métodos detectores.

Para isso, são analisados e discutidos os resultados da avaliação de desempenho com os trabalhos relacionados. Essa avaliação envolve a quantificação de resultados empíricos obtidos por cada método de detecção. A variação do cenário ocorre ao efetuar mudança no método de aprendizado e/ou gerador de dados. Assim sendo, são comparados os resultados entre diferentes métodos detectores de maneira quantitativa. A parte da análise qualitativa da avaliação é baseada na análise quantitativa dos resultados.

No tópico B.4, as Tabelas 26 e 27 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de Wilcoxon Signed-Rank) com as médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Nas Tabelas 28 e 29 apresentam os resultados dos testes das hipóteses estatísticas com as médias de MTR dos métodos testados em todos os conjuntos

de dados selecionados, bem como em suas classificações usando HT e NB. Seguidamente, nas Tabelas 30 e 31 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de *Wilcoxon Signed-Rank*) com as médias de MCC dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Observe que, para cada uma das três hipóteses H_0 , H_1 e H_2 testadas, na mesma linha o melhor resultado é mostrado em **negrito**. Essas tabelas apresentam a informação de qual método foi igual (empate), superior (vitória) ou inferior (derrota) estatisticamente a qualquer método avaliado. Na avaliação as derrotas são subtraídas das vitórias para obter o rank (das diferenças).

B.2 MOTIVAÇÃO

Assim como já descrito, o principal objetivo desta seção é avaliar o comportamento de técnicas de DMC quando utilizadas em cenários de fluxos contínuo de dados com MC. Para ser considerada aplicável neste cenário, uma técnica deve ser capaz de classificar as MC com boa acurácia e minimizar falsos alarmes de detecção. Um fator a ser considerado ao discutir-se a viabilidade de um método de detecção é sua eficiência em termos de três critérios de avaliação: (1) acurácia da classificação, (2) MTR e (3) MCC.

B.3 RESULTADOS E DISCUSSÃO

As Tabelas 20 e 21 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, vitórias, derrotas e média dos postos usando HT e NB, respectivamente. Os cenários avaliados nessas tabelas incluem três tamanhos de base de dados com simulações de mudanças abruptas e graduais, bem como em conjuntos de dados do mundo real. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos que obtiveram os melhores resultados com HT e NB foram o FHDDM e FSDD, respectivamente. O recurso do mapa de calor mostra uma maior concentração de vermelho nas linhas dos métodos ECDD, EDDM, SEED, STEPDP, SeqDrift1 e SeqDrift2 com o HT, já com o NB foram DDM, ECDD, EDDM, STEPDP, SeqDrift1 e SeqDrift2.

A Figura 33 (a) apresenta o resumo dos resultados com HT e NB, apresentados no tópico B.4 nas Tabelas 26 e 27. Em termos das hipóteses testadas, o FHDDM com HT forneceu o melhor desempenho com 9 vitórias e nenhuma derrota. Já com o NB os dois melhores foram o FSDD e RDDM, ambos com 10 vitórias e também sem derrotas. Por outro lado, os métodos FHDDMS, FPDD, FSDD, FTDD, HDDMA e RDDM foram iguais estatisticamente ao FHDDM com HT. No caso do NB, os métodos FHDDM, FHDDMS, FSDD, HDDMA e o HDDMW foram iguais estatisticamente ao FSDD e RDDM. O FPDD, apesar de ter obtido um resultado positivo com o HT (7), não conseguiu manter o mesmo padrão de desempenho com o NB, alcançando uma diferença de -1 . O método com maior número de derrotas usando tanto o HT quanto o NB foi o ECDD, sendo inferior a todos os

Tabela 20 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _o | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|--------------|-------|-------|-------|-------|--------|-------|-------|-------|-------------------|-------------------|-------|-------|-------|-------|-------|-------|
| A-20k-Agr | 85,68 | 85,71 | 83,50 | 86,92 | 86,95 | 86,20 | 86,44 | 86,20 | 86,67 | 86,92 | 86,80 | 86,74 | 86,95 | 84,83 | 85,65 | 86,99 |
| A-20k-LED | 69,68 | 67,93 | 68,63 | 69,98 | 70,06 | 68,55 | 68,48 | 68,54 | 70,02 | 69,91 | 70,12 | 55,99 | 64,95 | 67,83 | 63,26 | 68,24 |
| A-20k-Mix | 89,26 | 89,41 | 89,00 | 90,75 | 90,75 | 90,82 | 90,82 | 90,82 | 90,49 | 90,82 | 90,76 | 90,30 | 90,73 | 88,00 | 88,22 | 90,81 |
| A-20k-RBF | 73,98 | 73,29 | 73,47 | 74,83 | 74,78 | 74,72 | 74,94 | 74,81 | 74,79 | 74,73 | 74,77 | 74,71 | 74,68 | 74,26 | 74,37 | 74,96 |
| A-20k-Sine | 87,85 | 88,26 | 88,09 | 89,73 | 89,73 | 89,76 | 89,76 | 89,76 | 89,67 | 89,80 | 89,55 | 89,02 | 89,62 | 87,15 | 87,22 | 89,77 |
| A-20k-Wav | 78,46 | 78,52 | 78,26 | 79,32 | 79,36 | 79,34 | 79,38 | 79,37 | 79,29 | 79,36 | 79,31 | 79,27 | 79,45 | 78,95 | 78,89 | 79,39 |
| A-130k-Agr | 91,61 | 86,81 | 89,78 | 91,50 | 91,39 | 92,02 | 92,03 | 92,02 | 91,74 | 90,85 | 91,72 | 88,55 | 89,47 | 91,78 | 90,62 | 91,90 |
| A-130k-LED | 70,28 | 69,02 | 70,51 | 71,45 | 71,42 | 71,02 | 70,74 | 71,06 | 71,47 | 71,29 | 71,52 | 56,11 | 66,55 | 70,02 | 64,31 | 70,26 |
| A-130k-Mix | 92,35 | 89,85 | 90,93 | 92,60 | 92,60 | 92,58 | 92,58 | 92,58 | 92,57 | 92,62 | 92,29 | 91,72 | 91,22 | 92,11 | 92,08 | 92,59 |
| A-130k-RBF | 79,01 | 73,40 | 77,75 | 77,99 | 77,47 | 79,61 | 79,59 | 79,61 | 79,27 | 76,67 | 79,35 | 75,81 | 75,99 | 79,42 | 79,07 | 79,03 |
| A-130k-Sine | 93,03 | 88,36 | 90,95 | 93,34 | 93,34 | 93,34 | 93,34 | 93,34 | 93,28 | 93,32 | 92,99 | 90,41 | 90,72 | 92,87 | 92,80 | 93,32 |
| A-130k-Wav | 79,05 | 79,13 | 78,86 | 79,62 | 79,76 | 79,57 | 79,60 | 79,59 | 79,57 | 79,87 | 79,72 | 79,71 | 79,98 | 79,44 | 79,51 | 79,62 |
| A-890k-Agr | 95,18 | 86,97 | 93,92 | 95,16 | 95,12 | 95,26 | 95,26 | 95,26 | 95,14 | 94,91 | 93,92 | 88,96 | 90,51 | 95,22 | 93,75 | 95,19 |
| A-890k-LED | 70,66 | 69,29 | 71,08 | 71,71 | 71,68 | 71,72 | 71,25 | 71,70 | 71,71 | 71,58 | 71,74 | 56,23 | 66,95 | 72,23 | 64,46 | 70,63 |
| A-890k-Mix | 94,88 | 89,93 | 93,72 | 94,97 | 94,97 | 94,95 | 94,95 | 94,95 | 94,95 | 94,97 | 93,89 | 92,33 | 91,30 | 94,88 | 94,87 | 94,96 |
| A-890k-RBF | 85,07 | 73,40 | 83,69 | 83,55 | 81,80 | 85,18 | 85,09 | 85,16 | 84,37 | 79,46 | 83,28 | 76,01 | 76,89 | 85,16 | 84,34 | 83,83 |
| A-890k-Sine | 96,90 | 88,39 | 96,00 | 96,97 | 96,97 | 96,96 | 96,96 | 96,96 | 96,91 | 96,96 | 95,67 | 90,60 | 91,69 | 96,89 | 96,77 | 96,94 |
| A-890k-Wav | 81,87 | 79,18 | 81,57 | 81,22 | 80,07 | 82,19 | 81,97 | 82,17 | 81,76 | 79,97 | 80,42 | 79,94 | 80,04 | 82,22 | 82,20 | 81,13 |
| G-20k-Agr | 84,69 | 84,37 | 82,75 | 85,17 | 85,47 | 83,89 | 84,32 | 83,92 | 85,30 | 85,28 | 85,48 | 85,04 | 85,21 | 82,87 | 84,79 | 83,44 |
| G-20k-LED | 69,19 | 67,61 | 68,63 | 69,65 | 69,66 | 68,40 | 68,17 | 68,49 | 69,54 | 69,51 | 69,71 | 55,79 | 64,49 | 67,58 | 62,64 | 68,04 |
| G-20k-Mix | 87,50 | 86,74 | 87,69 | 87,77 | 87,42 | 87,18 | 87,25 | 87,22 | 87,48 | 87,54 | 87,59 | 86,79 | 87,29 | 87,78 | 87,41 | 87,45 |
| G-20k-RBF | 73,33 | 72,50 | 73,22 | 73,91 | 73,72 | 73,70 | 74,04 | 73,78 | 74,13 | 73,79 | 73,98 | 73,85 | 73,67 | 73,60 | 74,14 | 73,14 |
| G-20k-Sine | 86,85 | 86,03 | 86,73 | 86,99 | 86,75 | 86,44 | 86,53 | 86,47 | 86,74 | 86,84 | 86,89 | 86,11 | 86,39 | 85,96 | 86,64 | 86,73 |
| G-20k-Wav | 78,32 | 78,14 | 78,20 | 78,79 | 78,89 | 78,17 | 78,60 | 78,40 | 78,66 | 78,89 | 79,03 | 78,79 | 78,91 | 78,19 | 78,70 | 78,53 |
| G-130k-Agr | 91,61 | 86,59 | 89,77 | 91,19 | 91,13 | 91,72 | 91,74 | 91,72 | 91,51 | 90,58 | 91,57 | 88,30 | 89,14 | 91,65 | 90,53 | 90,99 |
| G-130k-LED | 69,89 | 68,96 | 70,54 | 71,39 | 71,35 | 70,95 | 70,76 | 70,98 | 71,36 | 71,23 | 71,45 | 56,15 | 66,37 | 70,01 | 64,31 | 70,16 |
| G-130k-Mix | 92,11 | 89,42 | 91,02 | 92,14 | 92,08 | 92,03 | 92,03 | 92,03 | 92,06 | 92,11 | 92,07 | 91,15 | 90,68 | 92,07 | 91,94 | 92,05 |
| G-130k-RBF | 79,09 | 73,27 | 77,66 | 77,68 | 77,24 | 79,32 | 79,38 | 79,41 | 79,15 | 76,49 | 79,37 | 75,62 | 76,02 | 79,31 | 79,02 | 78,49 |
| G-130k-Sine | 92,78 | 88,02 | 91,15 | 92,84 | 92,82 | 92,77 | 92,77 | 92,77 | 92,77 | 92,84 | 92,75 | 89,87 | 90,32 | 92,72 | 92,67 | 92,83 |
| G-130k-Wav | 79,05 | 78,68 | 78,93 | 79,68 | 79,65 | 79,75 | 79,63 | 79,84 | 79,59 | 79,70 | 79,75 | 79,59 | 79,76 | 79,67 | 79,58 | 79,43 |
| G-890k-Agr | 95,19 | 86,94 | 93,91 | 95,10 | 95,08 | 95,23 | 95,23 | 95,23 | 95,10 | 94,86 | 93,89 | 88,91 | 90,48 | 95,21 | 93,74 | 94,99 |
| G-890k-LED | 70,74 | 69,28 | 71,11 | 71,70 | 71,68 | 71,70 | 71,23 | 71,69 | 71,72 | 71,57 | 71,74 | 56,25 | 66,97 | 72,22 | 64,51 | 70,60 |
| G-890k-Mix | 94,87 | 89,86 | 93,73 | 94,90 | 94,88 | 94,86 | 94,85 | 94,86 | 94,88 | 94,88 | 93,86 | 92,24 | 91,22 | 94,87 | 94,86 | 94,87 |
| G-890k-RBF | 85,09 | 73,38 | 83,67 | 83,23 | 81,64 | 85,11 | 85,07 | 85,09 | 84,30 | 79,51 | 83,22 | 76,01 | 76,87 | 85,15 | 84,51 | 83,71 |
| G-890k-Sine | 96,86 | 88,34 | 95,98 | 96,89 | 96,87 | 96,85 | 96,84 | 96,84 | 96,85 | 96,88 | 95,73 | 90,55 | 91,56 | 96,89 | 96,72 | 96,86 |
| G-890k-Wav | 81,62 | 78,78 | 81,57 | 81,04 | 79,97 | 82,02 | 81,64 | 82,00 | 81,66 | 79,89 | 80,36 | 79,84 | 79,87 | 82,03 | 82,01 | 81,01 |
| Airlines | 65,43 | 64,20 | 65,62 | 65,73 | 65,81 | 65,80 | 65,71 | 65,77 | 66,00 | 65,43 | 66,54 | 66,26 | 65,81 | 66,07 | 66,15 | 65,93 |
| Connect4 | 74,70 | 75,00 | 74,76 | 75,21 | 75,14 | 74,41 | 74,88 | 74,57 | 75,14 | 75,21 | 75,18 | 74,83 | 75,20 | 73,97 | 74,35 | 75,18 |
| Outdoor | 59,02 | 58,84 | 58,80 | 60,09 | 60,25 | 59,66 | 58,46 | 59,21 | 59,96 | 60,85 | 58,63 | 59,09 | 59,74 | 56,97 | 57,95 | 60,12 |
| Rialto | 39,93 | 31,10 | 46,44 | 41,75 | 47,29 | 36,62 | 37,10 | 33,45 | 44,24 | 45,28 | 44,52 | 47,01 | 40,26 | 39,96 | 47,64 | 38,77 |
| SpamData | 92,11 | 92,60 | 91,96 | 92,57 | 92,92 | 92,18 | 92,18 | 92,18 | 92,76 | 92,46 | 92,30 | 92,85 | 92,21 | 91,44 | 91,48 | 92,43 |
| WineRed | 52,97 | 52,71 | 54,35 | 52,07 | 52,07 | 53,15 | 52,71 | 53,17 | 52,82 | 52,50 | 52,76 | 53,18 | 53,26 | 51,72 | 51,72 | 51,34 |
| WineWhite | 45,13 | 45,13 | 48,54 | 46,09 | 46,10 | 45,21 | 44,44 | 44,44 | 45,15 | 46,71 | 44,32 | 44,32 | 47,62 | 45,11 | 44,59 | 45,02 |
| Vitórias | 5 | 0 | 3 | 9 | 6 | 7 | 7 | 7 | 8 | 5 | 5 | 1 | 2 | 4 | 2 | 5 |
| Derrotas | 5 | 15 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 14 | 12 | 5 | 10 | 2 |
| Média postos | 9,03 | 14,08 | 11,07 | 5,34 | 6,13 | 6,88 | 7,08 | 6,66 | 6,19 | 6,42 | 6,84 | 12,05 | 10,56 | 8,99 | 10,85 | 7,84 |

Fonte: O AUTOR (2020).

métodos avaliados com o HT e NB. Além disso, os detectores EDDM e SeqDrift1 também foram inferiores a todos os métodos como NB.

Os melhores resultados foram obtidos pelos métodos FHDDM, FHDDMS, FPDD, FTDD, FSDD, HDDMA, HDDMW, RDDM e WSTD, todos com avaliação acima de zero, como pode ser visto na Figura 33 (b). Apesar disso, vale a pena destacar que o desempenho

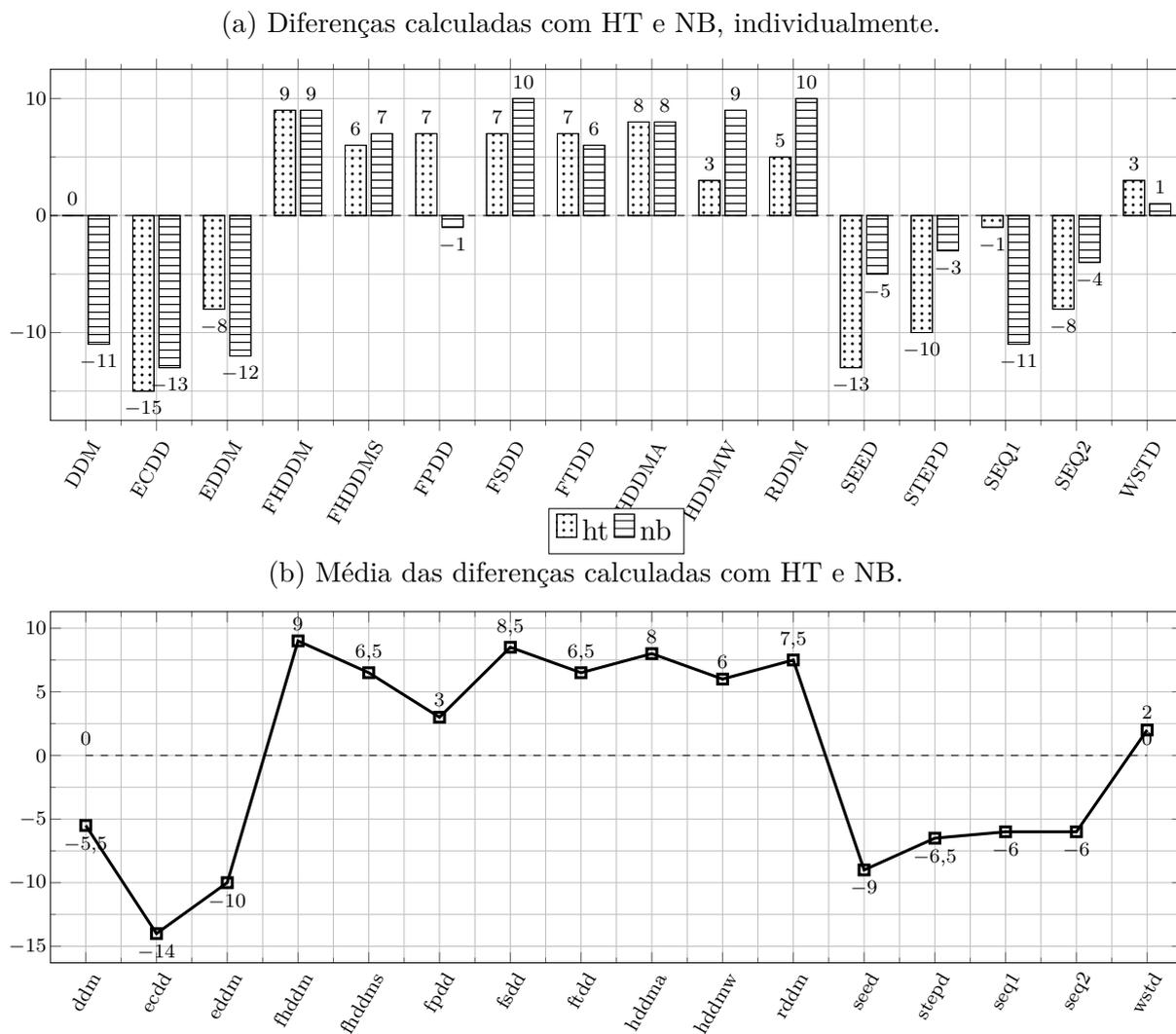
Tabela 21 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|--------------|-------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------------|-------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,07 | 85,57 | 84,42 | 86,95 | 86,99 | 85,71 | 86,83 | 85,71 | 86,86 | 86,94 | 86,89 | 86,72 | 87,07 | 84,08 | 85,48 | 86,99 |
| A-20k-LED | 70,21 | 67,99 | 69,01 | 71,04 | 71,08 | 66,15 | 67,97 | 66,10 | 71,40 | 70,56 | 71,83 | 56,84 | 65,57 | 65,23 | 63,71 | 68,06 |
| A-20k-Mix | 90,36 | 89,45 | 89,83 | 91,11 | 91,11 | 91,19 | 91,19 | 91,19 | 91,10 | 91,19 | 91,03 | 90,67 | 90,98 | 87,97 | 87,92 | 91,19 |
| A-20k-RBF | 72,69 | 73,45 | 73,26 | 74,88 | 74,91 | 74,76 | 74,89 | 74,83 | 74,77 | 74,90 | 74,72 | 74,79 | 74,77 | 74,00 | 74,21 | 74,84 |
| A-20k-Sine | 84,90 | 88,31 | 87,96 | 89,15 | 89,15 | 89,17 | 89,18 | 89,18 | 89,04 | 89,20 | 88,99 | 88,57 | 89,12 | 86,49 | 86,42 | 89,19 |
| A-20k-Wav | 78,42 | 78,37 | 77,90 | 79,70 | 79,78 | 79,43 | 79,69 | 79,55 | 79,46 | 79,81 | 79,63 | 79,58 | 79,72 | 78,91 | 79,21 | 79,66 |
| A-130k-Agr | 88,30 | 86,60 | 86,25 | 88,71 | 88,70 | 88,90 | 88,91 | 88,90 | 88,75 | 88,64 | 88,77 | 88,79 | 88,42 | 88,64 | 88,67 | 88,89 |
| A-130k-LED | 70,61 | 69,07 | 71,36 | 72,63 | 72,53 | 70,97 | 72,62 | 70,99 | 73,44 | 72,14 | 73,48 | 56,83 | 69,61 | 67,16 | 70,69 | 72,03 |
| A-130k-Mix | 91,07 | 89,87 | 89,99 | 91,93 | 91,93 | 91,94 | 91,94 | 91,94 | 91,84 | 91,94 | 91,84 | 91,84 | 91,55 | 91,45 | 91,44 | 91,94 |
| A-130k-RBF | 72,70 | 73,49 | 73,80 | 75,26 | 75,24 | 75,28 | 75,31 | 75,28 | 75,25 | 75,18 | 75,22 | 75,32 | 75,02 | 75,14 | 75,22 | 75,21 |
| A-130k-Sine | 84,19 | 88,39 | 86,78 | 89,48 | 89,48 | 89,49 | 89,49 | 89,49 | 89,37 | 89,49 | 89,37 | 89,35 | 89,33 | 89,07 | 89,05 | 89,49 |
| A-130k-Wav | 78,77 | 78,73 | 78,49 | 80,33 | 80,32 | 80,21 | 80,33 | 80,30 | 80,22 | 80,30 | 80,21 | 80,28 | 80,09 | 80,12 | 80,19 | 80,29 |
| A-890k-Agr | 88,22 | 86,79 | 86,71 | 89,10 | 89,07 | 89,37 | 89,36 | 89,37 | 89,30 | 89,01 | 89,25 | 89,30 | 88,63 | 89,34 | 89,34 | 89,29 |
| A-890k-LED | 70,43 | 69,30 | 71,74 | 73,02 | 72,89 | 73,71 | 73,82 | 73,70 | 73,87 | 72,49 | 73,84 | 57,62 | 70,10 | 70,62 | 73,48 | 73,36 |
| A-890k-Mix | 90,26 | 89,95 | 88,70 | 92,07 | 92,07 | 92,08 | 92,08 | 92,08 | 92,03 | 92,08 | 92,02 | 92,04 | 91,65 | 92,00 | 92,00 | 92,08 |
| A-890k-RBF | 72,34 | 73,49 | 73,58 | 75,36 | 75,31 | 75,44 | 75,46 | 75,46 | 75,43 | 75,23 | 75,40 | 75,45 | 75,10 | 75,40 | 75,46 | 75,40 |
| A-890k-Sine | 82,61 | 88,42 | 86,12 | 89,50 | 89,50 | 89,50 | 89,50 | 89,50 | 89,45 | 89,50 | 89,54 | 89,46 | 89,41 | 89,44 | 89,44 | 89,50 |
| A-890k-Wav | 78,58 | 78,80 | 78,42 | 80,38 | 80,36 | 80,40 | 80,40 | 80,40 | 80,35 | 80,33 | 80,37 | 80,38 | 80,16 | 80,37 | 80,38 | 80,39 |
| G-20k-Agr | 84,66 | 82,91 | 83,70 | 84,07 | 84,90 | 80,47 | 82,95 | 80,48 | 84,79 | 84,70 | 85,07 | 84,26 | 84,61 | 78,68 | 83,97 | 79,54 |
| G-20k-LED | 69,40 | 67,62 | 68,96 | 70,40 | 70,54 | 64,87 | 67,11 | 64,83 | 71,07 | 70,09 | 71,63 | 56,51 | 64,89 | 64,00 | 63,90 | 65,23 |
| G-20k-Mix | 87,82 | 86,89 | 87,98 | 88,18 | 87,85 | 87,64 | 87,67 | 87,65 | 87,81 | 87,99 | 88,00 | 86,97 | 87,48 | 88,01 | 87,70 | 87,86 |
| G-20k-RBF | 73,00 | 72,67 | 73,07 | 73,79 | 73,82 | 73,01 | 73,75 | 73,16 | 73,91 | 73,90 | 73,90 | 73,89 | 73,77 | 72,89 | 73,84 | 69,78 |
| G-20k-Sine | 86,96 | 86,18 | 86,71 | 87,11 | 86,90 | 86,55 | 86,55 | 86,56 | 86,96 | 87,03 | 87,00 | 85,80 | 86,32 | 85,82 | 86,33 | 86,75 |
| G-20k-Wav | 78,26 | 78,15 | 77,92 | 78,95 | 79,09 | 78,88 | 78,93 | 78,92 | 78,97 | 79,10 | 79,16 | 79,09 | 79,10 | 78,76 | 78,95 | 78,47 |
| G-130k-Agr | 87,93 | 86,32 | 86,22 | 88,31 | 88,32 | 88,35 | 88,39 | 88,36 | 88,45 | 88,25 | 88,50 | 88,37 | 88,01 | 88,15 | 88,43 | 86,62 |
| G-130k-LED | 70,95 | 69,01 | 71,37 | 72,54 | 72,44 | 70,36 | 72,11 | 70,35 | 73,41 | 72,06 | 73,47 | 56,98 | 69,53 | 66,03 | 70,33 | 71,17 |
| G-130k-Mix | 91,40 | 89,48 | 90,72 | 91,47 | 91,40 | 91,39 | 91,40 | 91,39 | 91,42 | 91,43 | 91,45 | 91,26 | 91,01 | 91,44 | 91,32 | 91,44 |
| G-130k-RBF | 72,58 | 73,36 | 73,86 | 75,06 | 75,03 | 75,07 | 75,07 | 75,06 | 75,11 | 74,99 | 75,12 | 75,16 | 74,84 | 75,02 | 75,15 | 73,78 |
| G-130k-Sine | 88,71 | 88,08 | 87,82 | 89,31 | 89,31 | 89,17 | 89,19 | 89,18 | 89,28 | 89,32 | 89,31 | 88,85 | 88,90 | 89,00 | 88,99 | 89,25 |
| G-130k-Wav | 78,71 | 78,70 | 78,47 | 80,11 | 80,14 | 79,81 | 80,08 | 80,02 | 80,09 | 80,11 | 80,08 | 80,15 | 79,99 | 79,78 | 80,13 | 79,64 |
| G-890k-Agr | 88,59 | 86,75 | 86,73 | 89,02 | 88,96 | 89,27 | 89,26 | 89,27 | 89,23 | 88,91 | 89,18 | 89,25 | 88,57 | 89,25 | 89,30 | 88,95 |
| G-890k-LED | 70,32 | 69,30 | 71,69 | 73,01 | 72,88 | 73,63 | 73,77 | 73,63 | 73,87 | 72,48 | 73,84 | 57,48 | 70,08 | 69,21 | 73,48 | 73,19 |
| G-890k-Mix | 91,54 | 89,89 | 88,34 | 92,00 | 91,99 | 91,99 | 92,00 | 91,99 | 91,99 | 92,00 | 91,95 | 91,96 | 91,57 | 92,00 | 91,99 | 92,00 |
| G-890k-RBF | 72,47 | 73,48 | 73,86 | 75,34 | 75,29 | 75,41 | 75,42 | 75,42 | 75,42 | 75,20 | 75,38 | 75,43 | 75,08 | 75,38 | 75,44 | 75,18 |
| G-890k-Sine | 84,24 | 88,38 | 86,96 | 89,50 | 89,50 | 89,48 | 89,47 | 89,48 | 89,45 | 89,50 | 89,53 | 89,39 | 89,35 | 89,43 | 89,43 | 89,47 |
| G-890k-Wav | 78,51 | 78,79 | 78,46 | 80,35 | 80,33 | 80,35 | 80,35 | 80,35 | 80,34 | 80,31 | 80,34 | 80,36 | 80,14 | 80,34 | 80,36 | 80,28 |
| Airlines | 65,35 | 63,66 | 65,18 | 65,82 | 65,83 | 66,72 | 66,91 | 66,76 | 67,23 | 65,34 | 67,50 | 66,81 | 65,73 | 65,67 | 66,59 | 66,84 |
| Connect4 | 74,47 | 75,05 | 74,80 | 75,23 | 75,22 | 74,20 | 74,65 | 74,23 | 74,95 | 75,23 | 75,23 | 74,76 | 75,14 | 73,11 | 74,17 | 75,12 |
| Outdoor | 59,49 | 60,71 | 59,28 | 60,59 | 60,63 | 59,77 | 60,23 | 59,80 | 60,11 | 61,97 | 59,85 | 59,26 | 61,60 | 57,26 | 58,04 | 61,03 |
| Rialto | 36,63 | 21,59 | 44,62 | 41,24 | 47,47 | 26,55 | 28,12 | 27,52 | 43,30 | 45,69 | 44,88 | 47,01 | 40,55 | 39,95 | 47,49 | 27,26 |
| SpamData | 89,34 | 88,39 | 90,89 | 91,51 | 91,73 | 91,35 | 91,35 | 91,35 | 90,96 | 91,68 | 91,39 | 91,62 | 91,42 | 90,92 | 89,00 | 91,80 |
| WineRed | 47,70 | 47,70 | 49,08 | 47,46 | 47,46 | 47,70 | 47,85 | 47,70 | 47,63 | 47,54 | 46,92 | 47,52 | 47,87 | 47,70 | 47,70 | 47,85 |
| WineWhite | 42,82 | 42,82 | 45,25 | 46,53 | 46,90 | 44,96 | 45,48 | 44,88 | 43,16 | 47,08 | 43,41 | 44,41 | 48,13 | 42,82 | 42,82 | 45,40 |
| Vitórias | 1 | 0 | 0 | 9 | 7 | 5 | 10 | 8 | 8 | 9 | 10 | 3 | 4 | 0 | 4 | 5 |
| Derrotas | 12 | 15 | 15 | 0 | 0 | 6 | 0 | 2 | 0 | 0 | 0 | 8 | 7 | 15 | 8 | 4 |
| Média Postos | 12,60 | 13,77 | 12,63 | 5,90 | 6,10 | 7,78 | 5,49 | 6,92 | 6,44 | 5,88 | 5,63 | 8,81 | 10,23 | 11,19 | 9,35 | 7,28 |

Fonte: O AUTOR (2020).

do FHDDM foi sólido na maioria das situações e líder em vitórias com HT e vice líder com NB. A Figura 33 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB, logo é possível observar que o FHDDM obteve a maior média (9,0) no desempenho acumulado, seguido pelo FSDD com 8,5, HDDMA 8,0, RDDM 7,5, FHDDMS e FTDD com 6,5, HDDMW 6,0, FPDD 3,0, WSTD com 2,0

Figura 33 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados.



Fonte: O AUTOR (2020).

e os demais ficaram com média negativa em número de vitórias.

As Tabelas 22 e 23 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos que obtiveram os melhores resultados com HT e NB foram o HDDMW e HDDMA, respectivamente. O recurso do mapa de calor mostra uma maior concentração de vermelho nas linhas dos métodos DDM, ECDD, EDDM, RDDM, SEED, STEPD, SeqDrift1 e SeqDrift2 com o HT, já com o NB foram DDM, ECDD, EDDM, STEPD, SeqDrift1 e SeqDrift2.

A Figura 34 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice B nas Tabelas 28 e 29. Em termos das hipóteses testas, o FSDD e HDDMW

Tabela 22 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEDP | SEQ1 | SEQ2 | WSTD |
|--------------|----------------|-------------|--------------|--------------|--------------|-----------------|-----------------|-----------------|-------------------|-------------------|---------------|-------------|--------------|--------|---------------|----------------|
| A-20k-Agr | 17,49 | 24,16 | 2,42 | 104,44 | 112,43 | 49,29 | 64,39 | 49,07 | 63,50 | 129,13 | 57,86 | 62,31 | 101,34 | 9,49 | 16,75 | 128,64 |
| A-20k-LED | 2,77 | 16,45 | 1,87 | 7,63 | 12,16 | 1,79 | 1,32 | 1,66 | 2,74 | 18,22 | 2,32 | 0,92 | 3,19 | 1,09 | 0,79 | 1,59 |
| A-20k-Mix | 112,34 | 290,17 | 33,15 | 310,41 | 310,41 | 340,77 | 340,77 | 340,77 | 257,89 | 439,48 | 267,77 | 87,83 | 427,78 | 20,00 | 17,33 | 360,21 |
| A-20k-RBF | 4,30 | 13,16 | 1,42 | 32,20 | 41,43 | 2,27 | 7,39 | 4,95 | 14,85 | 49,64 | 18,65 | 21,37 | 36,17 | 2,09 | 6,89 | 11,56 |
| A-20k-Sine | 49,48 | 218,76 | 19,06 | 283,61 | 283,61 | 285,17 | 285,17 | 285,17 | 252,81 | 361,46 | 129,05 | 86,11 | 337,71 | 20,00 | 17,16 | 314,69 |
| A-20k-Wav | 1,27 | 16,55 | 0,73 | 1,47 | 3,65 | 0,78 | 1,08 | 0,84 | 1,20 | 7,19 | 2,16 | 1,23 | 27,00 | 0,42 | 0,87 | 1,15 |
| A-130k-Agr | 65,51 | 34,37 | 4,35 | 732,18 | 805,05 | 961,42 | 1069,78 | 1011,47 | 702,40 | 818,32 | 169,33 | 196,52 | 232,80 | 126,75 | 99,48 | 1177,06 |
| A-130k-LED | 5,61 | 14,37 | 1,16 | 23,39 | 27,43 | 1,70 | 1,45 | 2,12 | 16,49 | 29,12 | 14,57 | 1,07 | 2,35 | 1,39 | 0,48 | 2,86 |
| A-130k-Mix | 199,11 | 372,06 | 8,35 | 1993,70 | 1993,70 | 1970,80 | 1970,80 | 1970,80 | 2497,95 | 2751,53 | 354,40 | 311,55 | 493,87 | 129,35 | 108,65 | 2288,86 |
| A-130k-RBF | 28,32 | 18,43 | 0,57 | 207,39 | 231,56 | 131,72 | 286,43 | 243,09 | 218,18 | 184,27 | 66,06 | 60,93 | 86,44 | 57,88 | 105,79 | 309,66 |
| A-130k-Sine | 152,53 | 161,79 | 5,85 | 1772,46 | 1772,46 | 2011,33 | 2011,33 | 2011,33 | 2568,86 | 2191,48 | 321,19 | 183,58 | 623,55 | 128,74 | 108,74 | 2142,92 |
| A-130k-Wav | 2,84 | 19,57 | 0,23 | 63,78 | 75,75 | 1,17 | 23,80 | 8,42 | 21,85 | 34,37 | 21,70 | 52,00 | 43,56 | 2,20 | 44,96 | 18,04 |
| A-890k-Agr | 228,85 | 25,95 | 3,02 | 5475,15 | 5637,86 | 9212,94 | 9069,74 | 9114,79 | 2518,23 | 6194,77 | 267,74 | 1172,84 | 211,37 | 881,74 | 643,33 | 8827,81 |
| A-890k-LED | 2,92 | 20,79 | 0,47 | 25,54 | 27,96 | 2,72 | 2,34 | 1,63 | 15,16 | 33,81 | 15,53 | 1,08 | 4,77 | 4,16 | 0,72 | 5,03 |
| A-890k-Mix | 601,94 | 369,92 | 7,48 | 13538,84 | 13538,84 | 16451,68 | 16451,68 | 16451,68 | 16981,12 | 17859,12 | 1007,84 | 388,50 | 443,92 | 890,00 | 735,65 | 16339,46 |
| A-890k-RBF | 126,59 | 12,10 | 0,57 | 1310,04 | 1129,41 | 2306,12 | 2075,84 | 2412,69 | 1099,40 | 493,44 | 197,84 | 47,11 | 77,15 | 753,53 | 623,35 | 1560,00 |
| A-890k-Sine | 635,25 | 198,46 | 12,13 | 11830,04 | 11830,04 | 16724,42 | 16724,42 | 16724,42 | 13992,72 | 14334,66 | 1250,35 | 180,16 | 654,16 | 885,55 | 710,79 | 15476,77 |
| A-890k-Wav | 25,64 | 19,49 | 0,00 | 218,15 | 62,42 | 108,74 | 449,29 | 298,90 | 556,74 | 64,63 | 24,23 | 58,07 | 55,62 | 238,57 | 619,19 | 222,57 |
| G-20k-Agr | 3,88 | 0,87 | 3,42 | 15,47 | 4,59 | 1,59 | 4,08 | 1,64 | 7,43 | 15,11 | 4,24 | 6,17 | 3,11 | 1,86 | 1,68 | 4,55 |
| G-20k-LED | 1,64 | 2,62 | 3,18 | 2,29 | 2,59 | 1,47 | 1,27 | 1,53 | 2,05 | 3,06 | 2,02 | 0,93 | 1,91 | 1,15 | 0,54 | 1,30 |
| G-20k-Mix | 2,20 | 1,19 | 34,62 | 0,95 | 0,91 | 5,01 | 5,49 | 4,98 | 2,68 | 0,92 | 1,32 | 6,70 | 9,60 | 0,48 | 1,25 | 5,98 |
| G-20k-RBF | 4,21 | 2,24 | 1,88 | 6,86 | 6,77 | 1,03 | 1,55 | 1,20 | 7,24 | 3,83 | 9,59 | 6,54 | 3,71 | 0,98 | 1,87 | 0,77 |
| G-20k-Sine | 3,92 | 0,94 | 13,99 | 1,32 | 0,89 | 2,17 | 2,17 | 2,08 | 1,20 | 0,92 | 2,67 | 7,25 | 9,81 | 1,09 | 0,33 | 4,41 |
| G-20k-Wav | 0,40 | 1,70 | 0,86 | 0,64 | 1,14 | 0,18 | 0,39 | 0,28 | 0,47 | 1,33 | 1,67 | 1,09 | 1,68 | 0,21 | 0,31 | 0,41 |
| G-130k-Agr | 65,79 | 0,50 | 4,55 | 129,63 | 43,49 | 70,01 | 107,01 | 96,21 | 99,82 | 69,43 | 153,23 | 26,64 | 6,88 | 9,79 | 0,72 | 90,16 |
| G-130k-LED | 2,97 | 2,22 | 1,21 | 8,38 | 10,88 | 1,55 | 1,73 | 1,71 | 4,93 | 7,35 | 7,92 | 1,35 | 1,30 | 1,27 | 0,40 | 1,59 |
| G-130k-Mix | 1037,63 | 0,94 | 11,50 | 0,87 | 0,99 | 28,77 | 32,21 | 29,52 | 45,98 | 5,02 | 95,13 | 28,81 | 12,92 | 0,30 | 11,31 | 36,70 |
| G-130k-RBF | 27,43 | 1,48 | 0,55 | 42,27 | 15,66 | 30,63 | 62,54 | 48,59 | 86,37 | 18,93 | 60,66 | 22,26 | 7,46 | 26,19 | 52,21 | 2,80 |
| G-130k-Sine | 23,05 | 0,49 | 10,87 | 1,50 | 0,79 | 8,27 | 20,95 | 8,42 | 0,87 | 0,85 | 212,37 | 14,26 | 18,24 | 0,73 | 0,77 | 13,15 |
| G-130k-Wav | 2,36 | 1,56 | 0,10 | 3,40 | 2,23 | 0,27 | 6,20 | 1,04 | 9,29 | 2,63 | 9,19 | 9,63 | 1,59 | 0,48 | 1,95 | 0,84 |
| G-890k-Agr | 270,20 | 0,53 | 2,69 | 1091,11 | 1,58 | 661,60 | 661,56 | 661,55 | 538,99 | 1574,17 | 142,82 | 151,64 | 1,66 | 1,47 | 0,52 | 1198,95 |
| G-890k-LED | 2,97 | 2,08 | 0,48 | 6,91 | 9,01 | 2,12 | 1,22 | 1,65 | 5,38 | 5,38 | 11,40 | 1,13 | 2,31 | 4,51 | 0,44 | 2,34 |
| G-890k-Mix | 918,06 | 1,38 | 7,52 | 173,78 | 0,80 | 303,00 | 294,47 | 308,53 | 0,95 | 57,85 | 116,91 | 31,28 | 13,84 | 18,20 | 182,57 | 381,50 |
| G-890k-RBF | 124,60 | 1,76 | 0,60 | 241,91 | 16,95 | 588,63 | 526,61 | 573,54 | 289,71 | 59,49 | 136,56 | 22,72 | 5,28 | 370,50 | 62,93 | 87,67 |
| G-890k-Sine | 1062,80 | 1,61 | 13,26 | 154,03 | 0,78 | 14,31 | 14,31 | 14,32 | 0,85 | 92,76 | 119,52 | 13,33 | 22,47 | 31,97 | 7,68 | 216,39 |
| G-890k-Wav | 19,87 | 1,44 | 0,00 | 17,08 | 6,10 | 27,96 | 44,76 | 47,32 | 145,44 | 5,19 | 11,67 | 10,22 | 1,29 | 90,51 | 121,14 | 1,70 |
| Vitórias | 5 | 0 | 3 | 9 | 6 | 7 | 7 | 7 | 8 | 5 | 5 | 1 | 2 | 4 | 2 | 5 |
| Derrotas | 5 | 15 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 14 | 12 | 5 | 10 | 2 |
| Média Postos | 9,03 | 11,25 | 12,94 | 5,83 | 7,42 | 8,25 | 6,53 | 7,50 | 6,17 | 5,53 | 7,00 | 9,69 | 8,47 | 11,92 | 11,86 | 6,61 |

Fonte: O AUTOR (2020).

com HT forneceram os melhores desempenhos, ambos com 9 vitórias e nenhuma derrota, já com NB o FSDD obteve 12 vitórias e também nenhuma derrota. Entretanto, os métodos FHDDM, FPDD, FTDD, HDDMA e WSTD foram iguais estatisticamente aos métodos FSDD e HDDMW com HT e com NB apenas os métodos HDDMA, HDDMW e WSTD foram iguais estatisticamente ao FSDD. O SEED obteve um resultado positivo (3) com NB em comparação a sua versão com HT (-7), houve uma grande perda na eficiência usando HT. O método de maior número de derrotas usando HT foi o EDDM com 15 derrotas (todos os métodos avaliados foram superiores estatisticamente a ele) e repete-se com NB, tratando-se da ineficiência ele foi em termos absolutos e em termos estatísticos o pior método avaliado na métrica MTR.

Nota-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos métodos FHDDM, FHDDMS, FPDD, FSDD (10, 5), FTDD, HDDMA, HDDMW, RDDM e WSTD, todos com avaliação acima de zero, mostrados na Figura 34 (b). Apesar disso, vale a pena destacar que o desempenho do FSDD foi sólido na maioria das situações e

Tabela 23 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _h | HDDM _w | RDDM | SEED | STEDP | SEQ1 | SEQ2 | WSTD |
|--------------|---------------|-------------|--------------|--------------|--------------|---------------|-----------------|----------------|-------------------|-------------------|--------------|---------------|---------------|--------|--------|---------------|
| A-20k-Agr | 20,39 | 19,60 | 6,74 | 116,51 | 121,04 | 67,39 | 94,07 | 66,48 | 87,95 | 132,61 | 52,64 | 59,96 | 109,18 | 15,06 | 17,19 | 140,66 |
| A-20k-LED | 5,18 | 16,68 | 1,20 | 44,97 | 57,04 | 0,46 | 0,81 | 0,45 | 19,59 | 93,59 | 15,06 | 0,96 | 3,68 | 0,35 | 0,94 | 11,99 |
| A-20k-Mix | 92,02 | 292,17 | 34,39 | 306,08 | 306,08 | 345,14 | 345,14 | 345,14 | 459,62 | 426,00 | 183,33 | 104,31 | 605,20 | 20,00 | 18,11 | 374,36 |
| A-20k-RBF | 6,40 | 10,87 | 1,67 | 25,85 | 39,94 | 3,57 | 10,34 | 3,62 | 20,44 | 36,93 | 19,67 | 27,50 | 32,40 | 2,51 | 7,98 | 10,96 |
| A-20k-Sine | 36,50 | 218,02 | 21,76 | 284,58 | 284,58 | 283,03 | 293,51 | 288,32 | 221,29 | 363,52 | 122,49 | 95,55 | 323,64 | 20,00 | 18,03 | 299,98 |
| A-20k-Wav | 2,30 | 15,53 | 0,55 | 51,84 | 88,85 | 7,09 | 34,64 | 22,15 | 27,17 | 114,05 | 25,68 | 36,58 | 89,90 | 0,33 | 13,02 | 35,21 |
| A-130k-Agr | 52,91 | 32,89 | 2,94 | 531,93 | 573,75 | 585,69 | 721,14 | 594,73 | 289,91 | 776,10 | 142,37 | 331,64 | 174,00 | 117,47 | 105,15 | 835,10 |
| A-130k-LED | 8,08 | 14,24 | 1,40 | 59,68 | 69,91 | 4,83 | 26,87 | 4,86 | 87,89 | 54,67 | 49,55 | 1,17 | 9,86 | 0,95 | 5,05 | 9,86 |
| A-130k-Mix | 206,23 | 374,26 | 10,81 | 2037,77 | 2037,77 | 2277,70 | 2277,70 | 2277,70 | 1375,12 | 2836,13 | 514,13 | 635,31 | 728,17 | 130,00 | 116,01 | 2468,60 |
| A-130k-RBF | 9,43 | 14,48 | 0,75 | 89,45 | 133,69 | 19,37 | 136,42 | 56,55 | 89,95 | 98,74 | 55,05 | 187,37 | 58,37 | 6,24 | 90,12 | 115,92 |
| A-130k-Sine | 45,76 | 161,18 | 5,29 | 1894,08 | 1871,53 | 1919,41 | 1987,90 | 1953,63 | 642,88 | 2453,78 | 252,86 | 671,32 | 497,73 | 130,00 | 115,38 | 1992,14 |
| A-130k-Wav | 11,36 | 17,91 | 0,25 | 447,25 | 406,65 | 166,23 | 496,92 | 367,31 | 199,48 | 414,47 | 77,73 | 268,04 | 96,65 | 84,48 | 112,64 | 442,11 |
| A-890k-Agr | 92,18 | 21,07 | 1,79 | 4038,41 | 4215,80 | 4124,63 | 4630,12 | 4142,22 | 765,94 | 5061,79 | 210,27 | 1523,04 | 159,73 | 842,04 | 702,83 | 4777,01 |
| A-890k-LED | 8,18 | 20,54 | 1,46 | 76,17 | 87,26 | 125,22 | 218,26 | 112,85 | 339,96 | 59,45 | 75,76 | 1,20 | 10,26 | 29,27 | 24,34 | 16,19 |
| A-890k-Mix | 347,10 | 361,59 | 8,58 | 13633,88 | 13633,88 | 14959,61 | 14959,61 | 14959,61 | 4215,07 | 18776,21 | 940,79 | 2743,57 | 790,08 | 890,00 | 747,14 | 16327,63 |
| A-890k-RBF | 10,09 | 11,03 | 0,75 | 227,29 | 151,45 | 181,23 | 465,73 | 323,00 | 211,93 | 112,91 | 51,81 | 759,85 | 62,84 | 37,54 | 675,54 | 168,54 |
| A-890k-Sine | 42,93 | 196,49 | 2,93 | 11627,48 | 9440,65 | 12156,83 | 12724,30 | 12527,86 | 1368,98 | 12427,04 | 285,46 | 2873,47 | 503,47 | 890,00 | 775,03 | 11163,95 |
| A-890k-Wav | 7,62 | 10,30 | 0,27 | 1070,72 | 610,60 | 1633,53 | 1859,05 | 2051,98 | 695,58 | 520,56 | 86,27 | 1031,56 | 149,88 | 667,49 | 758,64 | 1160,33 |
| G-20k-Agr | 6,65 | 1,13 | 5,87 | 13,21 | 7,71 | 4,01 | 5,48 | 3,87 | 13,13 | 12,04 | 11,84 | 4,22 | 2,30 | 2,56 | 1,32 | 3,86 |
| G-20k-LED | 3,46 | 2,42 | 1,08 | 6,21 | 13,33 | 0,29 | 0,65 | 0,28 | 8,91 | 7,49 | 11,76 | 1,07 | 1,30 | 0,22 | 0,40 | 0,37 |
| G-20k-Mix | 0,90 | 0,72 | 41,79 | 1,06 | 0,84 | 3,48 | 3,53 | 3,48 | 0,88 | 0,97 | 8,14 | 7,01 | 12,85 | 0,51 | 1,16 | 5,24 |
| G-20k-RBF | 8,02 | 2,22 | 2,06 | 4,94 | 4,31 | 2,61 | 3,58 | 2,68 | 9,94 | 4,85 | 11,83 | 7,04 | 3,02 | 2,14 | 2,95 | 1,17 |
| G-20k-Sine | 4,26 | 0,85 | 25,21 | 1,45 | 0,90 | 3,19 | 2,30 | 3,72 | 0,89 | 0,83 | 3,57 | 8,47 | 8,87 | 1,19 | 0,39 | 5,58 |
| G-20k-Wav | 1,02 | 1,69 | 0,64 | 0,45 | 0,64 | 0,31 | 0,30 | 0,30 | 0,49 | 1,42 | 1,43 | 0,89 | 1,27 | 0,32 | 0,37 | 0,28 |
| G-130k-Agr | 49,15 | 0,89 | 3,54 | 76,86 | 27,87 | 59,45 | 65,85 | 57,85 | 132,99 | 67,27 | 129,43 | 26,53 | 1,43 | 10,30 | 0,91 | 42,13 |
| G-130k-LED | 8,84 | 1,95 | 1,42 | 13,38 | 16,97 | 1,28 | 9,89 | 1,19 | 71,71 | 9,36 | 51,60 | 1,10 | 3,67 | 0,97 | 2,49 | 1,35 |
| G-130k-Mix | 740,13 | 0,76 | 17,81 | 0,88 | 12,22 | 28,83 | 33,44 | 30,32 | 99,87 | 1,06 | 68,91 | 54,53 | 18,88 | 0,32 | 19,41 | 20,83 |
| G-130k-RBF | 10,42 | 1,53 | 0,73 | 17,44 | 14,96 | 14,64 | 24,40 | 24,94 | 51,30 | 11,40 | 46,85 | 34,25 | 4,00 | 4,61 | 24,77 | 4,00 |
| G-130k-Sine | 106,09 | 0,47 | 7,29 | 1,08 | 0,99 | 16,64 | 8,57 | 12,23 | 41,45 | 12,28 | 65,55 | 53,82 | 10,58 | 0,94 | 0,76 | 15,64 |
| G-130k-Wav | 7,94 | 1,56 | 0,26 | 64,74 | 33,03 | 21,33 | 76,30 | 48,06 | 62,82 | 22,11 | 64,89 | 34,68 | 1,51 | 19,86 | 8,86 | 17,99 |
| G-890k-Agr | 127,79 | 0,49 | 1,86 | 411,57 | 66,77 | 453,48 | 511,37 | 453,50 | 464,38 | 409,72 | 71,56 | 154,63 | 1,39 | 83,48 | 11,08 | 292,95 |
| G-890k-LED | 9,78 | 2,15 | 1,33 | 17,73 | 22,15 | 48,37 | 67,03 | 49,08 | 273,23 | 10,48 | 59,50 | 1,21 | 2,76 | 14,92 | 18,69 | 1,95 |
| G-890k-Mix | 565,96 | 1,44 | 10,56 | 0,81 | 0,83 | 247,61 | 223,79 | 220,70 | 1062,98 | 1,06 | 92,94 | 259,46 | 21,38 | 0,36 | 94,04 | 394,77 |
| G-890k-RBF | 20,33 | 1,65 | 0,65 | 30,89 | 9,41 | 128,12 | 131,11 | 162,65 | 193,52 | 9,97 | 43,46 | 140,88 | 2,86 | 49,51 | 86,75 | 15,40 |
| G-890k-Sine | 70,15 | 0,61 | 2,74 | 1,39 | 1,08 | 207,09 | 12,11 | 178,08 | 409,75 | 1,09 | 95,84 | 227,39 | 14,70 | 0,91 | 0,50 | 312,75 |
| G-890k-Wav | 9,49 | 1,47 | 0,26 | 124,52 | 10,88 | 421,03 | 137,51 | 403,37 | 365,24 | 31,74 | 42,07 | 88,30 | 2,27 | 332,89 | 39,51 | 6,30 |
| Vitórias | 1 | 1 | 0 | 8 | 6 | 7 | 12 | 11 | 7 | 7 | 5 | 6 | 2 | 1 | 1 | 7 |
| Derrotas | 10 | 11 | 15 | 2 | 3 | 2 | 0 | 1 | 0 | 0 | 6 | 3 | 9 | 10 | 10 | 0 |
| Média Postos | 10,28 | 12,53 | 13,19 | 6,56 | 7,31 | 7,67 | 5,42 | 6,94 | 5,11 | 5,94 | 7,25 | 7,50 | 9,28 | 12,69 | 11,14 | 7,19 |

Fonte: O AUTOR (2020).

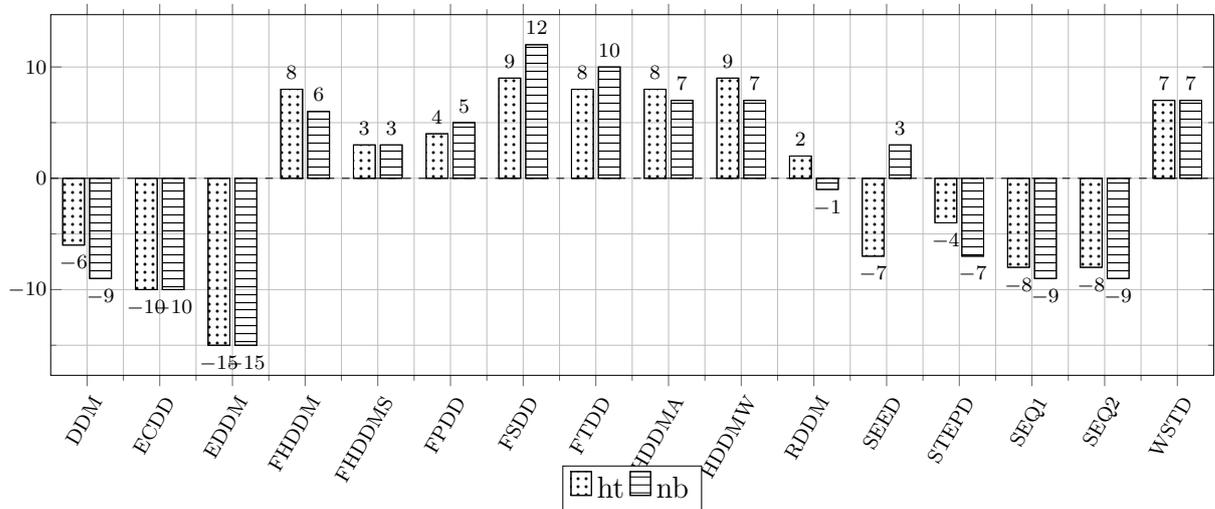
líder em vitórias com NB (12), seguido do FTDD com 11 e com HT foi juntamente lider com HDDMW, ambos com 9 vitórias. A Figura 34 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB, logo é possível observar que o FSDD obteve a maior média (10,5) no desempenho acumulado, seguido pelo FTDD com 9, HDDMW (8,0), HDDMA (7,5), FHDDM e WSTD com 7, FTDD (4,5), FHDDMS (3,0), RDDM com 0,5 e os demais ficaram com média negativa em número de vitórias.

As Tabelas 24 e 25 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os métodos que obtiveram os melhores resultados com HT e NB, foram o FHDDMS e FHDDM, respectivamente. O recurso do mapa de calor mostra uma maior concentração de vermelho nas linhas dos métodos DDM, ECDD, EDDM, SeqDrift1 e SeqDrift2 com o HT e NB.

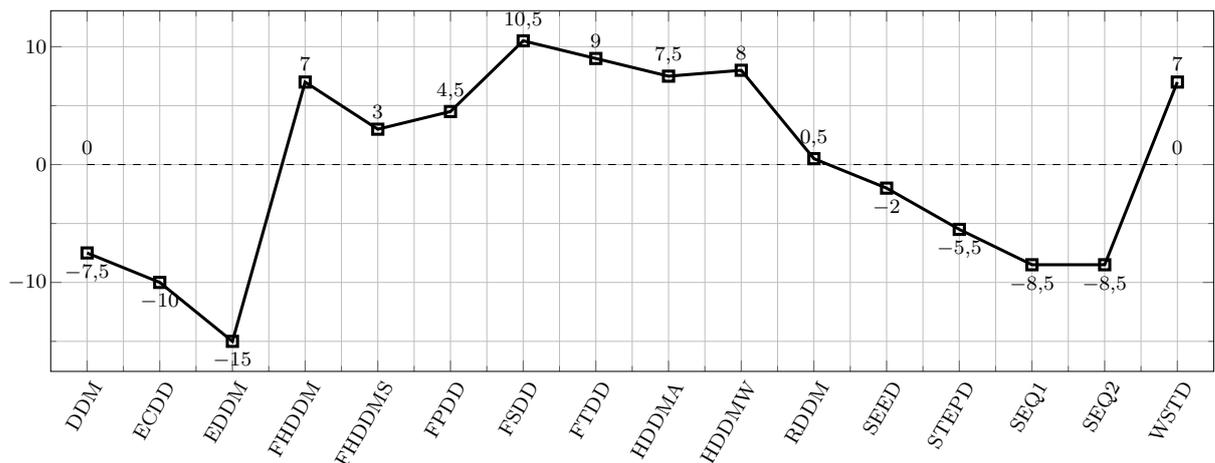
A Figura 35 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice B nas Tabelas 30 e 31. Em termos das hipóteses testadas, o FHDDM e FHDDMS com HT forneceram os melhores desempenhos, ambos com 10 vitórias e nenhuma derrota. Já com NB os métodos FHDDM, FHDDMS, FPDD, FSDD, FTDD e WSTD obtiveram

Figura 34 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



Fonte: O AUTOR (2020).

9 vitórias e também nenhuma derrota. Entretanto, os métodos FTDD, FSDD, FTDD e WSTD foram iguais estatisticamente aos métodos FHDDM e FHDDMS com HT e com NB foram iguais os métodos FHDDM, FHDDMS, FPDD, FSDD, FTDD e WSTD. O SEED obteve um resultado positivo (3) com NB em comparação a sua versão com HT (-2). Os métodos com maiores números de derrotas usando HT foram SeqDrift1 e SeqDrift2 com 14 derrotas (todos os outros métodos avaliados foram superiores estatisticamente a eles) e com NB o método SeqDrift1 obteve 15 derrotas. Assim sendo, tratando-se da ineficiência os métodos SeqDrift1 e SeqDrift2 foram, em termos absolutos e em termos estatísticos, os piores métodos avaliados.

Os melhores resultados foram obtidos pelos métodos FHDDM, FHDDMS, FPDD, FSDD, FTDD, WSTD, HDDMW, HDDMA e SEED, todos com avaliação acima de zero,

Tabela 24 – Médias calculadas do MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|--------------|-------------|-------|-------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------------|-------------|-------------|-------|-------|-------|-------------|
| A-20k-Agr | 0,33 | 0,28 | 0,18 | 0,88 | 0,93 | 0,79 | 0,83 | 0,79 | 0,81 | 0,86 | 0,73 | 0,78 | 0,45 | 0,05 | -0,01 | 0,93 |
| A-20k-LED | 0,01 | 0,24 | 0,06 | 0,23 | 0,32 | 0,03 | 0,03 | 0,02 | 0,13 | 0,37 | 0,03 | 0,05 | 0,16 | 0,12 | 0,11 | 0,12 |
| A-20k-Mix | 0,96 | 0,37 | 0,54 | 1,00 | 1,00 | 0,95 | 0,95 | 0,95 | 0,97 | 0,99 | 0,86 | 0,75 | 0,63 | -0,01 | 0,02 | 0,97 |
| A-20k-RBF | 0,07 | 0,25 | 0,07 | 0,62 | 0,66 | 0,56 | 0,66 | 0,61 | 0,47 | 0,61 | 0,36 | 0,58 | 0,53 | -0,01 | -0,01 | 0,72 |
| A-20k-Sine | 0,77 | 0,34 | 0,27 | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 1,00 | 1,00 | 0,85 | 0,71 | 0,69 | -0,01 | 0,01 | 0,98 |
| A-20k-Wav | 0,10 | 0,21 | 0,08 | 0,48 | 0,48 | 0,54 | 0,52 | 0,56 | 0,39 | 0,47 | 0,32 | 0,51 | 0,44 | 0,00 | -0,01 | 0,51 |
| A-130k-Agr | 0,13 | 0,11 | 0,07 | 0,80 | 0,86 | 0,93 | 0,94 | 0,93 | 0,87 | 0,70 | 0,22 | 0,34 | 0,23 | 0,00 | 0,00 | 0,94 |
| A-130k-LED | 0,00 | 0,09 | 0,00 | 0,11 | 0,15 | 0,02 | 0,01 | 0,02 | 0,07 | 0,16 | 0,01 | 0,01 | 0,04 | 0,04 | 0,06 | 0,05 |
| A-130k-Mix | 0,73 | 0,14 | 0,01 | 1,00 | 1,00 | 0,94 | 0,94 | 0,94 | 0,97 | 1,00 | 0,82 | 0,46 | 0,28 | 0,00 | 0,00 | 0,95 |
| A-130k-RBF | 0,00 | 0,11 | 0,00 | 0,45 | 0,46 | 0,68 | 0,70 | 0,69 | 0,53 | 0,34 | 0,00 | 0,29 | 0,28 | 0,00 | 0,00 | 0,66 |
| A-130k-Sine | 0,40 | 0,12 | 0,01 | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,96 | 0,99 | 0,73 | 0,35 | 0,34 | 0,00 | 0,00 | 0,98 |
| A-130k-Wav | 0,01 | 0,08 | 0,01 | 0,38 | 0,34 | 0,56 | 0,44 | 0,53 | 0,19 | 0,26 | 0,12 | 0,37 | 0,20 | 0,00 | 0,00 | 0,44 |
| A-890k-Agr | 0,00 | 0,04 | 0,00 | 0,84 | 0,85 | 0,98 | 0,97 | 0,98 | 0,69 | 0,69 | 0,13 | 0,14 | 0,09 | 0,00 | 0,00 | 0,94 |
| A-890k-LED | 0,00 | 0,04 | 0,00 | 0,05 | 0,05 | 0,01 | 0,01 | 0,00 | 0,03 | 0,07 | 0,01 | 0,00 | 0,02 | 0,00 | 0,01 | 0,03 |
| A-890k-Mix | 0,00 | 0,06 | 0,00 | 1,00 | 1,00 | 0,94 | 0,94 | 0,94 | 0,95 | 1,00 | 0,40 | 0,20 | 0,11 | 0,00 | 0,00 | 0,95 |
| A-890k-RBF | 0,00 | 0,04 | 0,00 | 0,33 | 0,26 | 0,72 | 0,70 | 0,74 | 0,42 | 0,15 | 0,05 | 0,11 | 0,11 | 0,00 | 0,00 | 0,49 |
| A-890k-Sine | 0,19 | 0,05 | 0,00 | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,97 | 0,99 | 0,37 | 0,14 | 0,14 | 0,00 | 0,00 | 0,97 |
| A-890k-Wav | 0,00 | 0,04 | 0,00 | 0,25 | 0,14 | 0,54 | 0,38 | 0,50 | 0,26 | 0,12 | 0,05 | 0,15 | 0,08 | 0,00 | 0,00 | 0,26 |
| Vitórias | 3 | 3 | 0 | 10 | 10 | 8 | 8 | 8 | 8 | 8 | 5 | 6 | 4 | 0 | 0 | 8 |
| Derrotas | 10 | 11 | 15 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 8 | 8 | 9 | 15 | 15 | 0 |
| Média Postos | 12,75 | 11,00 | 14,33 | 3,92 | 3,75 | 5,72 | 5,39 | 5,83 | 6,17 | 4,72 | 10,83 | 9,56 | 10,00 | 13,56 | 14,03 | 4,44 |

Fonte: O AUTOR (2020).

mostrados na Figura 35 (b). Apesar disso, vale a pena destacar que o desempenho dos métodos FHDDM e FHDDMS foi sólido na maioria das situações e líder em vitórias com HT (ambos com 10 vitórias), seguido dos métodos FPDD, FSDD, FTDD e WSTD com 8. Já com o NB os melhores detectores foram o FPDD, FSDD, FTDD e WSTD, ambos com 8 vitórias. A Figura 35 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que os métodos FHDDM e FHDDMS obtiveram as maiores médias (9,5) no desempenho acumulado, seguidos pelos métodos FPDD, FSDD, FTDD e WSTD com 8,5, HDDMW 6, HDDMA 3 e SEED com 0,5. Os demais métodos ficaram com médias negativas em números de vitórias.

A Figura 36 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 36 (a) é feito o rank da média das diferenças. A Figura 36 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 36 (a) ou no da Figura 36 (b), o detector de maior valor absoluto foi o FSDD, ambos com média 9,17. Vale ressaltar que por meio do rank da média (Figura 36 (a)) é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários.

O FSDD foi o único método no qual em todas as avaliações não obteve derrota. Além

Tabela 25 – Médias calculadas do MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados.

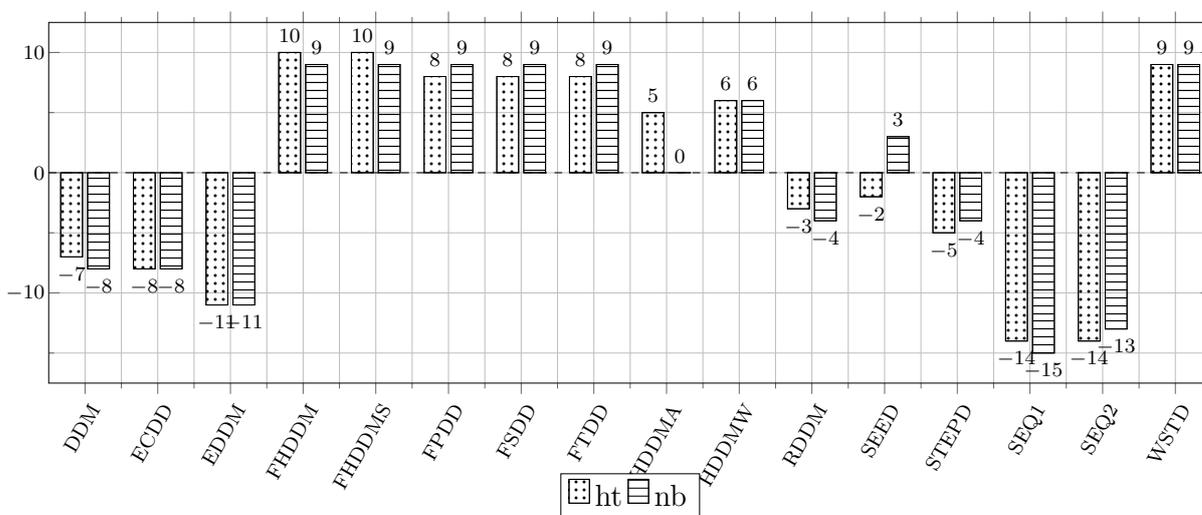
| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|--------------|-------|-------|-------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------------|-------|------|-------|-------|-------|-------------|
| A-20k-Agr | 0,26 | 0,27 | 0,20 | 0,91 | 0,93 | 0,78 | 0,87 | 0,77 | 0,84 | 0,86 | 0,71 | 0,76 | 0,43 | -0,01 | -0,01 | 0,96 |
| A-20k-LED | -0,01 | 0,24 | 0,02 | 0,49 | 0,59 | 0,08 | 0,10 | 0,07 | 0,17 | 0,52 | 0,01 | 0,05 | 0,17 | 0,01 | 0,08 | 0,24 |
| A-20k-Mix | 0,94 | 0,36 | 0,55 | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 1,00 | 1,00 | 0,97 | 0,87 | 0,78 | -0,01 | -0,01 | 1,00 |
| A-20k-RBF | 0,05 | 0,23 | 0,05 | 0,68 | 0,66 | 0,65 | 0,71 | 0,69 | 0,55 | 0,60 | 0,39 | 0,61 | 0,53 | -0,01 | -0,01 | 0,74 |
| A-20k-Sine | 0,67 | 0,33 | 0,32 | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 0,99 | 0,99 | 0,91 | 0,82 | 0,69 | -0,01 | -0,01 | 0,99 |
| A-20k-Wav | 0,04 | 0,20 | 0,03 | 0,71 | 0,85 | 0,47 | 0,65 | 0,57 | 0,48 | 0,85 | 0,33 | 0,66 | 0,62 | 0,00 | -0,01 | 0,67 |
| A-130k-Agr | 0,11 | 0,11 | 0,04 | 0,62 | 0,61 | 0,85 | 0,90 | 0,85 | 0,58 | 0,55 | 0,26 | 0,66 | 0,21 | 0,00 | 0,00 | 0,86 |
| A-130k-LED | 0,00 | 0,09 | 0,00 | 0,21 | 0,26 | 0,08 | 0,17 | 0,08 | 0,11 | 0,21 | 0,01 | 0,01 | 0,09 | 0,00 | 0,00 | 0,18 |
| A-130k-Mix | 0,47 | 0,14 | 0,02 | 1,00 | 1,00 | 0,99 | 0,99 | 0,99 | 0,95 | 1,00 | 0,90 | 0,76 | 0,33 | 0,00 | 0,00 | 1,00 |
| A-130k-RBF | 0,02 | 0,10 | 0,01 | 0,34 | 0,34 | 0,65 | 0,58 | 0,61 | 0,26 | 0,29 | 0,06 | 0,62 | 0,26 | 0,00 | 0,00 | 0,56 |
| A-130k-Sine | 0,19 | 0,12 | 0,04 | 0,99 | 0,97 | 0,99 | 0,98 | 0,99 | 0,72 | 0,95 | 0,52 | 0,77 | 0,32 | 0,00 | 0,00 | 0,97 |
| A-130k-Wav | 0,01 | 0,07 | 0,00 | 0,66 | 0,67 | 0,53 | 0,72 | 0,63 | 0,29 | 0,57 | 0,07 | 0,63 | 0,24 | 0,00 | 0,00 | 0,68 |
| A-890k-Agr | 0,00 | 0,04 | 0,00 | 0,29 | 0,28 | 0,84 | 0,83 | 0,84 | 0,26 | 0,24 | 0,15 | 0,50 | 0,08 | 0,00 | 0,00 | 0,55 |
| A-890k-LED | 0,00 | 0,04 | 0,00 | 0,10 | 0,11 | 0,05 | 0,12 | 0,05 | 0,05 | 0,09 | 0,01 | 0,00 | 0,03 | 0,00 | 0,00 | 0,08 |
| A-890k-Mix | 0,00 | 0,06 | 0,00 | 1,00 | 1,00 | 0,98 | 0,98 | 0,98 | 0,59 | 1,00 | 0,31 | 0,61 | 0,14 | 0,00 | 0,00 | 0,99 |
| A-890k-RBF | 0,00 | 0,04 | 0,00 | 0,14 | 0,13 | 0,47 | 0,36 | 0,46 | 0,12 | 0,11 | 0,03 | 0,40 | 0,10 | 0,00 | 0,00 | 0,24 |
| A-890k-Sine | 0,01 | 0,05 | 0,00 | 0,92 | 0,68 | 0,94 | 0,92 | 0,94 | 0,40 | 0,61 | 0,20 | 0,58 | 0,13 | 0,00 | 0,00 | 0,78 |
| A-890k-Wav | 0,00 | 0,02 | 0,00 | 0,34 | 0,29 | 0,52 | 0,47 | 0,57 | 0,12 | 0,21 | 0,04 | 0,48 | 0,09 | 0,00 | 0,00 | 0,39 |
| Vitórias | 3 | 3 | 2 | 9 | 9 | 9 | 9 | 9 | 7 | 8 | 5 | 7 | 5 | 0 | 1 | 9 |
| Derrotas | 11 | 11 | 13 | 0 | 0 | 0 | 0 | 0 | 7 | 2 | 9 | 4 | 9 | 15 | 14 | 0 |
| Média Postos | 13,22 | 11,28 | 14,67 | 3,83 | 3,89 | 5,11 | 4,11 | 4,94 | 7,72 | 5,53 | 10,72 | 7,56 | 10,06 | 14,33 | 15,06 | 3,97 |

Fonte: O AUTOR (2020).

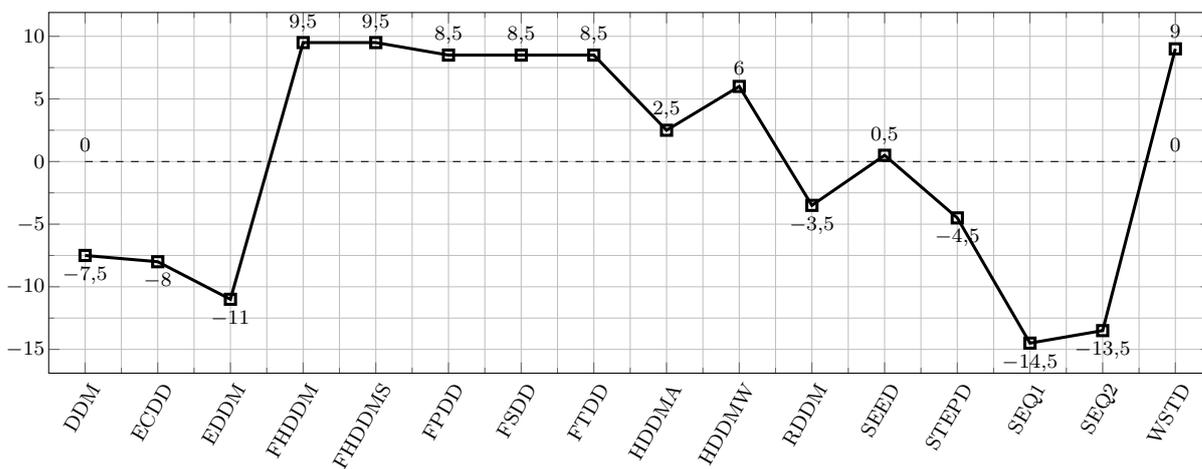
disso, ele também foi o método de melhor classificação ordenada por cálculo de média e por cálculo dos postos. Assim, o FSDD será utilizado como detector de referência, levando em consideração os critérios adotados nesta tese.

Figura 35 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados.

(a) Diferenças calculadas com HT e NB, individualmente.



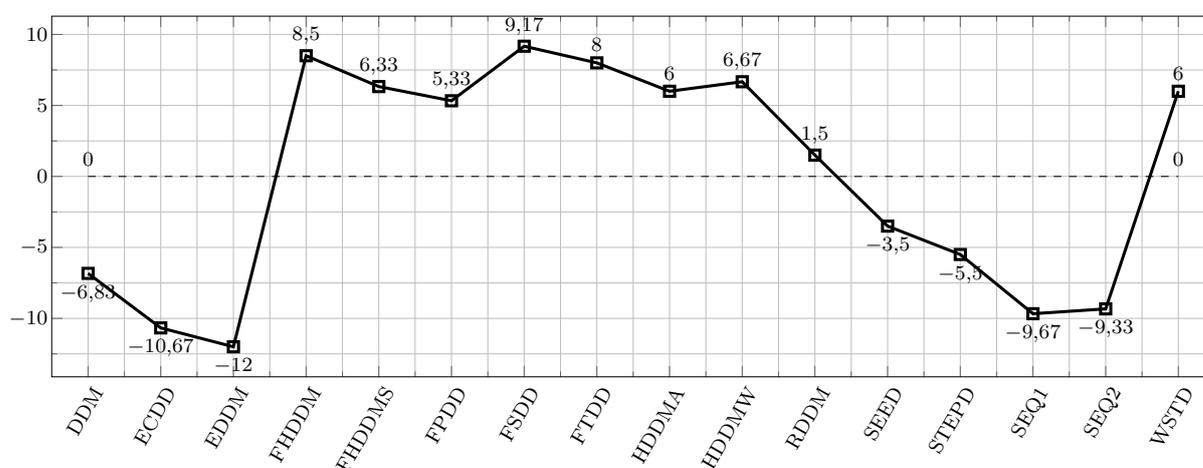
(b) Média das diferenças calculadas com HT e NB.



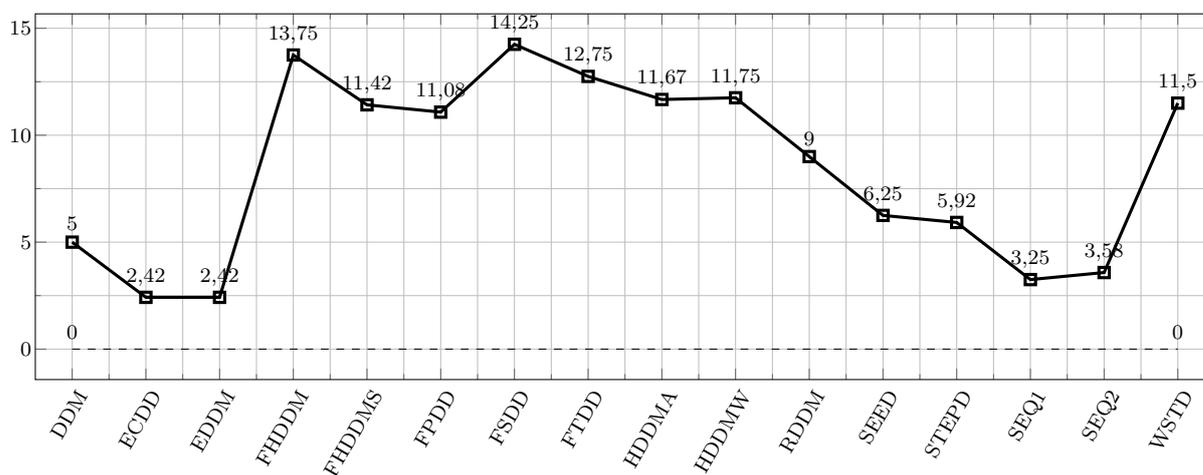
Fonte: O AUTOR (2020).

Figura 36 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando as comparações com os métodos de única análise dos trabalhos relacionados.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.



Fonte: O AUTOR (2020).

B.4 RESUMO DOS TESTES DE HIPÓTESES

Tabela 26 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₂ | H ₂ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ |
| ecdd | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ |
| eddm | H ₁ | H ₂ | * | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ | H ₀ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₀ | H ₀ | H ₂ |
| fhddms | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ |
| fpdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| fsdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₂ |
| hddmw | H ₀ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ |
| rddm | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ |
| seed | H ₁ | H ₂ | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ |
| stepd | H ₁ | H ₂ | H ₁ | H ₁ | H ₁ | H ₂ | * | H ₁ | H ₀ | H ₁ |
| seq1 | H ₀ | H ₂ | H ₂ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₂ | * | H ₀ | H ₀ |
| seq2 | H ₁ | H ₂ | H ₀ | H ₁ | H ₁ | H ₁ | H ₂ | H ₀ | H ₀ | * | H ₁ |
| wstd | H ₀ | H ₂ | H ₂ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ | * |
| $\sum H_0$ - Empates | 5 | 0 | 1 | 6 | 9 | 8 | 8 | 8 | 7 | 8 | 10 | 0 | 1 | 6 | 3 | 8 |
| $\sum H_1$ - Vitórias | 5 | 0 | 3 | 9 | 6 | 7 | 7 | 7 | 8 | 5 | 5 | 1 | 2 | 4 | 2 | 5 |
| $\sum H_2$ - Derrotas | 5 | 15 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 14 | 12 | 5 | 10 | 2 |
| $\sum H_1 - \sum H_2$ | 0 | -15 | -8 | 9 | 6 | 7 | 7 | 7 | 8 | 3 | 5 | -13 | -10 | -1 | -8 | 3 |
| Rank | 10,0 | 16,0 | 12,5 | 1,0 | 6,0 | 4,0 | 4,0 | 4,0 | 2,0 | 8,5 | 7,0 | 15,0 | 14,0 | 11,0 | 12,5 | 8,5 |

Fonte: O AUTOR (2020).

Tabela 27 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₂ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ |
| ecdd | H ₁ | * | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ |
| eddm | H ₀ | H ₀ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ |
| fhddms | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| fpdd | H ₂ | H ₂ | H ₂ | H ₁ | H ₀ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ |
| fsdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| hddmw | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ |
| rddm | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₂ |
| seed | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | * | H ₀ | H ₀ | H ₀ | H ₁ |
| stepd | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₂ | H ₀ | H ₀ |
| seq1 | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₁ | H ₁ |
| seq2 | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | * | H ₀ |
| wstd | H ₂ | H ₂ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | H ₂ | H ₀ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 2 | 2 | 3 | 6 | 8 | 4 | 5 | 5 | 7 | 6 | 5 | 4 | 4 | 4 | 3 | 6 |
| $\sum H_1$ - Vitórias | 1 | 0 | 0 | 9 | 7 | 5 | 10 | 8 | 8 | 9 | 10 | 3 | 4 | 0 | 4 | 5 |
| $\sum H_2$ - Derrotas | 12 | 13 | 12 | 0 | 0 | 6 | 0 | 2 | 0 | 0 | 0 | 8 | 7 | 11 | 8 | 4 |
| $\sum H_1 - \sum H_2$ | -11 | -13 | -12 | 9 | 7 | -1 | 10 | 6 | 8 | 9 | 10 | -5 | -3 | -11 | -4 | 1 |
| Rank | 13,5 | 16,0 | 15,0 | 3,5 | 6,0 | 9,0 | 1,5 | 7,0 | 5,0 | 3,5 | 1,5 | 12,0 | 10,0 | 13,5 | 11,0 | 8,0 |

Fonte: O AUTOR (2020).

Tabela 28 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ |
| ecdd | H ₀ | * | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₁ |
| eddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₀ | H ₂ | H ₀ |
| fhddms | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₁ |
| fpdd | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| fsdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₂ | H ₀ |
| hddmw | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ |
| rddm | H ₂ | H ₂ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | * | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| seed | H ₀ | H ₂ | H ₂ | H ₁ | H ₁ | H ₁ | * | H ₀ | H ₀ | H ₀ | H ₁ |
| stepd | H ₀ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₁ |
| seq1 | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | * | H ₀ | H ₁ |
| seq2 | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | * | H ₁ |
| wstd | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 7 | 3 | 0 | 7 | 4 | 9 | 6 | 7 | 7 | 6 | 5 | 4 | 7 | 5 | 5 | 8 |
| $\sum H_1$ - Vitórias | 1 | 1 | 0 | 8 | 7 | 5 | 9 | 8 | 8 | 9 | 6 | 2 | 2 | 1 | 1 | 7 |
| $\sum H_2$ - Derrotas | 7 | 11 | 15 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 4 | 9 | 6 | 9 | 9 | 0 |
| $\sum H_1 - \sum H_2$ | -6 | -10 | -15 | 8 | 3 | 4 | 9 | 8 | 8 | 9 | 2 | -7 | -4 | -8 | -8 | 7 |
| Rank | 11,0 | 15,0 | 16,0 | 4,0 | 8,0 | 7,0 | 1,5 | 4,0 | 4,0 | 1,5 | 9,0 | 12,0 | 10,0 | 13,5 | 13,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 29 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₁ |
| ecdd | H ₀ | * | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | H ₁ |
| eddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₀ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ |
| fhddms | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| fpdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| fsdd | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ |
| ftdd | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| hddmw | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| rddm | H ₂ | H ₂ | H ₂ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₁ |
| seed | H ₂ | H ₂ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | H ₀ | H ₀ | * | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| stepd | H ₀ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₀ | H ₁ |
| seq1 | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₀ | H ₁ |
| seq2 | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₀ | * | H ₁ |
| wstd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 4 | 3 | 0 | 5 | 6 | 6 | 3 | 3 | 8 | 8 | 4 | 6 | 4 | 4 | 4 | 8 |
| $\sum H_1$ - Vitórias | 1 | 1 | 0 | 8 | 6 | 7 | 12 | 11 | 7 | 7 | 5 | 6 | 2 | 1 | 1 | 7 |
| $\sum H_2$ - Derrotas | 10 | 11 | 15 | 2 | 3 | 2 | 0 | 1 | 0 | 0 | 6 | 3 | 9 | 10 | 10 | 0 |
| $\sum H_1 - \sum H_2$ | -9 | -10 | -15 | 6 | 3 | 5 | 12 | 10 | 7 | 7 | -1 | 3 | -7 | -9 | -9 | 7 |
| Rank | 13,0 | 15,0 | 16,0 | 6,0 | 8,5 | 7,0 | 1,0 | 2,0 | 4,0 | 4,0 | 10,0 | 8,5 | 11,0 | 13,0 | 13,0 | 4,0 |

Fonte: O AUTOR (2020).

Tabela 30 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₂ | H ₂ | H ₁ |
| ecdd | H ₀ | * | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ |
| eddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| fhddms | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ |
| fpdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ |
| fsdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₁ |
| hddmw | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ |
| rddm | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₁ |
| seed | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | * | H ₂ | H ₂ | H ₂ | H ₁ |
| stepd | H ₀ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₂ | H ₂ | H ₁ |
| seq1 | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | * | H ₀ | H ₁ |
| seq2 | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₁ |
| wstd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | * |
| $\sum H_0$ - Empates | 2 | 1 | 0 | 5 | 5 | 7 | 7 | 7 | 4 | 5 | 2 | 1 | 2 | 1 | 1 | 6 |
| $\sum H_1$ - Vitórias | 3 | 3 | 2 | 10 | 10 | 8 | 8 | 8 | 8 | 8 | 5 | 6 | 4 | 0 | 0 | 9 |
| $\sum H_2$ - Derrotas | 10 | 11 | 13 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 8 | 8 | 9 | 14 | 14 | 0 |
| $\sum H_1 - \sum H_2$ | -7 | -8 | -11 | 10 | 10 | 8 | 8 | 8 | 5 | 6 | -3 | -2 | -5 | -14 | -14 | 9 |
| Rank | 12,0 | 13,0 | 14,0 | 1,5 | 1,5 | 5,0 | 5,0 | 5,0 | 8,0 | 7,0 | 10,0 | 9,0 | 11,0 | 15,5 | 15,5 | 3,0 |

Fonte: O AUTOR (2020).

Tabela 31 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos de única análise dos trabalhos relacionados.

| Dataset | DDM | ECDD | EDDM | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | SEED | STEPD | SEQ1 | SEQ2 | WSTD |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ddm | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ |
| ecdd | H ₀ | * | H ₂ | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ |
| eddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₁ |
| FHDDM | H ₂ | H ₂ | H ₂ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| fhddms | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| fpdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ |
| fsdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₀ |
| hddma | H ₂ | H ₂ | H ₂ | H ₁ | * | H ₁ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₁ |
| hddmw | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ |
| rddm | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | * | H ₁ | H ₀ | H ₂ | H ₂ | H ₁ |
| seed | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₁ | H ₁ | H ₁ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₂ | H ₂ | H ₁ |
| stepd | H ₂ | H ₂ | H ₂ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₂ | H ₂ | H ₁ |
| seq1 | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₁ |
| seq2 | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₂ | * | H ₁ |
| wstd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | * |
| $\sum H_0$ - Empates | 1 | 1 | 0 | 6 | 6 | 6 | 6 | 6 | 1 | 5 | 1 | 4 | 1 | 0 | 0 | 6 |
| $\sum H_1$ - Vitórias | 3 | 3 | 2 | 9 | 9 | 9 | 9 | 9 | 7 | 8 | 5 | 7 | 5 | 0 | 1 | 9 |
| $\sum H_2$ - Derrotas | 11 | 11 | 13 | 0 | 0 | 0 | 0 | 0 | 7 | 2 | 9 | 4 | 9 | 15 | 14 | 0 |
| $\sum H_1 - \sum H_2$ | -8 | -8 | -11 | 9 | 9 | 9 | 9 | 9 | 0 | 6 | -4 | 3 | -4 | -15 | -13 | 9 |
| Rank | 12,5 | 12,5 | 14,0 | 3,5 | 3,5 | 3,5 | 3,5 | 3,5 | 9,0 | 7,0 | 10,5 | 8,0 | 10,5 | 16,0 | 15,0 | 3,5 |

Fonte: O AUTOR (2020).

B.5 CONSIDERAÇÕES FINAIS

Foi possível classificar o desempenho dos métodos avaliados em termos estatísticos usando teste de hipótese e média dos postos. Os resultados experimentais mostram que o detector com melhor média geral na classificação, considerando os três critérios de avaliação foi o FSDD, seguido do FHDDM e FTDD. Considerando a média geral apenas do critério de acurácia da classificação, o melhor foi o FHDDM, seguido do FSDD e HDDMA. Por fim, os resultados de desempenhos dos métodos foi calculado com base na metodologia de avaliação proposta nesta tese apresentada na seção 3 que classifica o desempenho com base nas diferenças estatística calculadas entre pares de tratamento do grupo de tratamentos.

A análise dos resultados indicou que alguns métodos bons e outros ruins. Considerando como parâmetros de bom e ruim os métodos de melhor e pior desempenho, a diferença entre os resultados passa dos 100% entre os valores absolutos alcançados em relação as médias obtidas pelos testes de hipóteses. O modo utilizado para comparação dos tratamentos apontou para a necessidade de se continuar fazendo este tipo de investigação de desempenho analítica, fundamentada em análise pareada dos métodos para formalização dos resultados.

APÊNDICE C – COMITÊS DA LITERATURA

Na seção anterior foram avaliados os métodos de DMC do estado da arte em MC. Os resultados experimentais mostram que o método FSDD obteve o melhor desempenho, considerando a média de desempenho nos critérios avaliados. No tópico C.3 são avaliados os desempenhos de quatro métodos de comitês de detectores: AD, ALHD, ALO (WOZNIAK et al., 2016) e *e*-Detector (DU et al., 2014). São utilizados quatro detectores de única análise como membros para formar os comitês: FHDDM, FSDD, FTDD e HDDMA. Os membros foram selecionados com base nos resultados da avaliação de desempenho do apêndice B e para isso foram adotados dois critérios de seleção de membros: melhor desempenho no rank (1) da média das diferenças dos três critérios de avaliação; e (2) da média das diferenças do critério de acurácia da classificação com HT e NB. Foram criados três comitês de referência: $C_1 = \{FHDDM, FSDD, FTDD\}$, $C_2 = \{FHDDM, FSDD, HDDMA\}$ e $C_3 = \{FHDDM, FSDD\}$. Resultados experimentais mostram que o *e*-Detector obteve os melhores resultados em comparação com o estado da arte usando o comitê C_2 . O restante desta seção está organizada da seguinte maneira: no tópico C.1 são apresentadas as considerações iniciais e explicações em relação aos objetivos esperados da avaliação; seguidamente, C.2 são apresentados os resultados e discussão; por fim, em C.4 apresenta as considerações finais sobre os resultados obtidos.

C.1 CONSIDERAÇÕES INICIAIS

Este tópico C trata de avaliar os comitês de detectores apresentados na seção de trabalhos relacionados, especificamente aos trabalhos de detecção de múltipla análise. A partir desta avaliação de desempenho, pretende-se identificar, classificar e ordenar os métodos detectores de melhor desempenho e selecioná-los para serem comparados com os métodos de comitês propostos nesta tese na seção 4. Assim sendo, pretende-se classificar o desempenho do estado da arte em métodos de comitês de detectores.

A partir dos resultados obtidos na seção B foram selecionados os detectores para compor os três comitês utilizados neste tópico C e na seção 4 para avaliar os métodos de comitês de detectores. Os membros dos comitês foram selecionados com base em três critérios de seleção: (1) em C_1 , foram selecionados os três melhores resultados no rank do cálculo das diferenças estatísticas, considerando as médias de resultados das métricas de acurácia da classificação, MCC e MTR; (2) em C_2 , foram selecionados os três melhores resultados no rank do cálculo dos postos apenas considerando a métrica de acurácia da classificação; e (3) em C_3 , foram selecionados os dois melhores métodos considerando a interseção entre os dois comitês C_1 e C_2 .

As comparações estatísticas serão feitas entre comitês de mesmo método e métrica.

Assim sendo, dado os conjuntos comitês $C = \{C_1, C_2 \text{ e } C_3\}$ e métodos $M = \{AD, ALHD, ALO \text{ e } e\text{-Detector}\}$ haverá comparações de cada método do conjunto M com seus respectivos comitês. O objetivo das comparações é identificar os comitês de melhor desempenho por método do conjunto M .

No tópico C.3, as Tabelas 38 e 39 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de *Wilcoxon Signed-Rank*) com as médias de acurácias dos métodos testados nos conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. As Tabelas 42 e 43 apresentam os resultados dos testes das hipóteses estatísticas com as médias de MCC dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Seguidamente, as Tabelas 40 e 41 apresentam os resultados dos testes das hipóteses estatísticas (testadas com o teste de *Wilcoxon Signed-Rank*) com as médias de MTR dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB. Visualmente foi proposto que as tabelas mostrem em **negrito** os melhores resultados, assim como o mapa de calor que identifica em vermelhos e verdes graduais os piores e melhores resultados, respectivamente. No entanto, observe que o mapa de calor e destaque em **negrito** apresentam apenas uma visão geral em relação a todos os métodos (linha completa). É importante ressaltar que as comparações visuais são ilustrativas, pois as avaliações são feitas somente em relação aos comitês de um mesmo método.

Essas tabelas, citadas no parágrafo anterior, apresentam a informação de qual método foi igual (empate), superior (vitória) ou inferior (derrota) estatisticamente a qualquer método avaliado. Na avaliação as derrotas são subtraídas das vitórias para obter o rank (das diferenças).

C.2 RESULTADOS E DISCUSSÃO

As Tabelas 32 e 33 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. O objetivo é identificar a melhor variação de comitê para cada método. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês que obtiveram os melhores resultados com o HT e NB foram ADC_3 , $ALHDC_3$, $ALOC_1$, $ALOC_2$ e $eDetC_2$. O recurso do mapa de calor mostra uma maior concentração de vermelho nas linhas dos métodos AD e ALHD com o HT e NB, implicando nas variações com piores resultados. Mais especificamente o $AD.C_2$ e $ALHDC_2$, tanto para HT quanto para NB, obtiveram os piores resultados em relação a seus respectivos comitês. Vale ressaltar, também, que os comitês dos métodos ALHD e ALO obtiveram melhores resultados nas mudanças do tipo abruptas do que nas graduais tanto para NB quanto para HT. Diferentemente do *e-Detector* que teve essa métrica

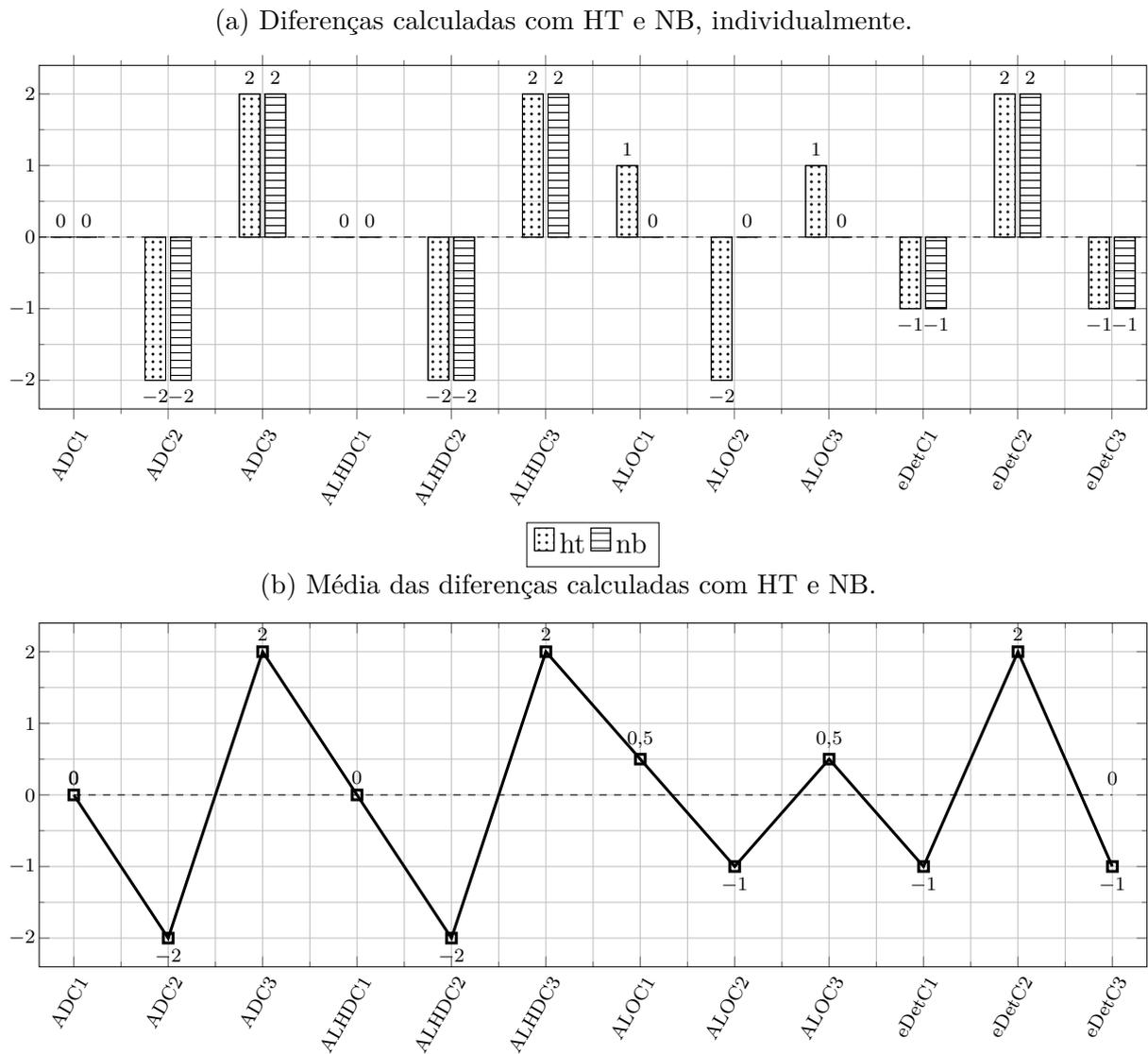
contrária. Avaliando os resultados das bases de dados reais, verificou-se melhores resultados para os comitês $ALHDC_3$, $ALOC_1$, $ALOC_2$, $ALOC_2$ e $eDetC_2$ para HT e NB.

Para visualizar o comportamento e comparativo dos métodos foram traçadas as diferenças entre vitórias e derrotas (consulte a Figura 37). Para cada método são apresentados as diferenças individuais com HT e NB além de serem calculadas as médias considerando os resultados de ambos classificadores. Algumas observações sobre o comportamento dos métodos podem ser deduzidas das Figuras 37a e 37b. Primeiro, as parametrizações ADC_3 , $ALHDC_3$ e $ALOC_2$ dos métodos AD, ALHD e ALO tendem a obter os melhores resultados. Por exemplo, as médias de acurácias entre eles foram as mais altas das três parametrizações avaliadas. Já as versões $AD.C_2$, $ALHDC_2$ e $ALOC_2$ foram as piores. Segundo, o *e-Detector* obteve os melhores resultados com a parametrização $eDetC_2$. Já com as outras duas os resultados foram iguais. Terceiro, nota-se que os métodos *e-Detector* e ALO são bem parecidos pois a diferença entre eles está apenas no uso do warning (não utilizada pelo ALO). Essa característica do uso do warning pelo *e-Detector*, quando comparado com os resultados do ALO, foi responsável pela piora nos resultados do *e-Detector* nas bases de tamanhos menores e na medida que os conceitos aumentaram o uso do *warning* favoreceu o *e-Detector*.

É possível notar que os melhores resultados nos conjuntos de dados avaliados foram obtidos pelos comitês ADC_3 , $ALHDC_3$, $ALOC_1$, $ALOC_2$ e $eDetC_2$, todos com diferenças calculadas acima de zero, como pode ser visto na Figura 37 (b). A Figura 37 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo é possível observar que os comitês ADC_3 , $ALHDC_3$, $ALOC_1$, $ALOC_2$ e $eDetC_2$ obtiveram a maior média entre as variações de comitês de cada método, considerando o desempenho acumulado, assim como $AD.C_2$ e $ALHDC_2$ obtiveram a pior média. Vale ressaltar que apenas cinco dos doze comitês avaliados obtiveram rank com valor positivo e diferente de zero. Por fim, comparando os métodos entre si, assumindo absolutos das médias dos experimentos com as bases abruptas, graduais e reais, com HT e NB, as parametrizações mais eficientes foram $ALOC_1$, $ALHDC_3$, $ALOC_2$ e $ALOC_2$, com isso é possível deduzir que o ALO foi o método mais eficaz independente da parametrização utilizada.

As Tabelas 34 e 35 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês ADC_3 , $ALHDC_3$ e $eDetC_2$ obtiveram os melhores resultados com o HT e NB. Os comitês $ALOC_1$, $ALOC_2$ e $ALOC_2$ foram iguais estatisticamente em relação a essas médias dos . Assim como nos resultados das médias das acurácias, o mapa de calor mostra uma grande concentração de vermelho nas linhas dos métodos AD, $ALHDC_1$ e $ALHDC_2$ com o HT e NB, implicando nas variações com piores resultados. Destaca-se também o $AD.C_2$ para o método AD e o $ALHDC_2$ para o método ALHD, tanto para

Figura 37 – Resumo dos ranks das acurácias com HT e NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados.



Fonte: O AUTOR (2020).

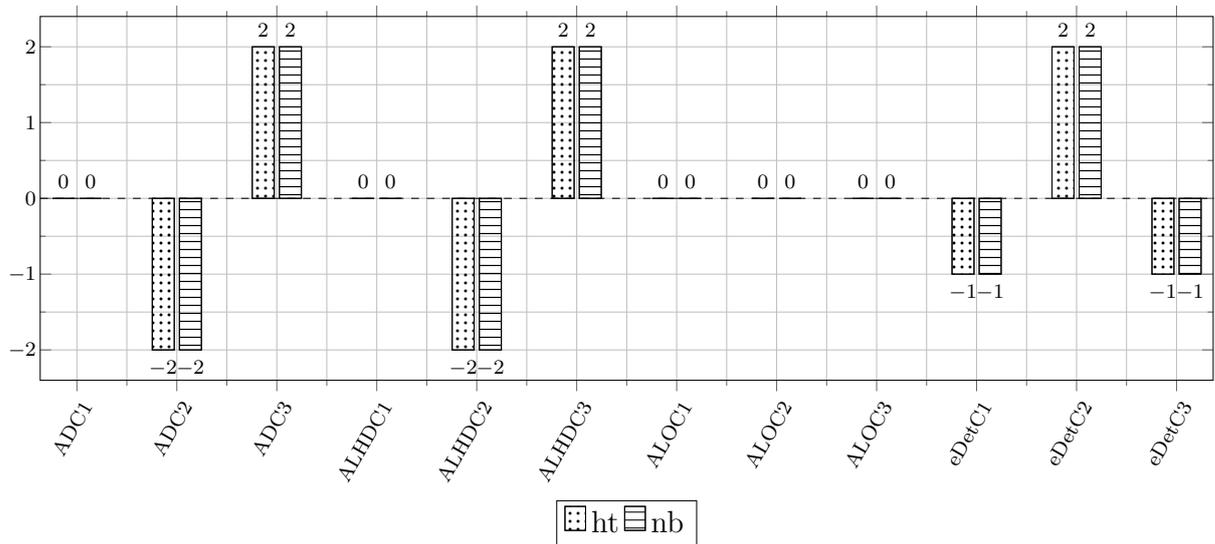
HT quanto para NB, obtendo os piores resultados. Neste caso, os resultados dos comitês do método ALO obteve melhor desempenho nas mudanças do tipo abruptas, tanto para NB quanto para HT. O oposto ocorreu com o método *e-Detector*, alcançando melhores resultados nas mudanças graduais.

A Figura 38 (a) traz o resumo dos resultados com HT e NB apresentados nas Tabelas 40 e 41. Em termos das hipóteses testadas, os comitês ADC₃, ALHDC₃ e eDetC₂ com o HT e NB, forneceram os melhores desempenhos e foram estatisticamente superiores aos demais comitês de mesmo método. Por outro lado, o AD.C₂ e o ALHDC₂ obtiveram os piores desempenhos. Tratando-se do ALO, os comitês ALOC₁, ALOC₂ e ALOC₂ foram iguais estatisticamente em relação ao rank dos MTRs com o HT e NB. Em relação ao rank dos

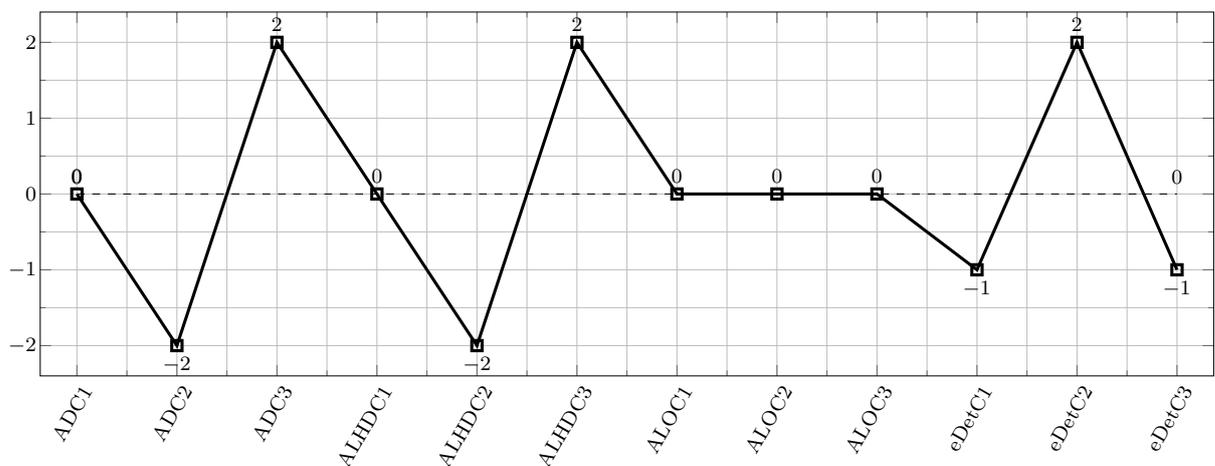
, somente três dos doze comitês avaliados conquistaram um rank com valores acima de zero.

Figura 38 – Resumo dos ranks dos MTRs com HT e NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



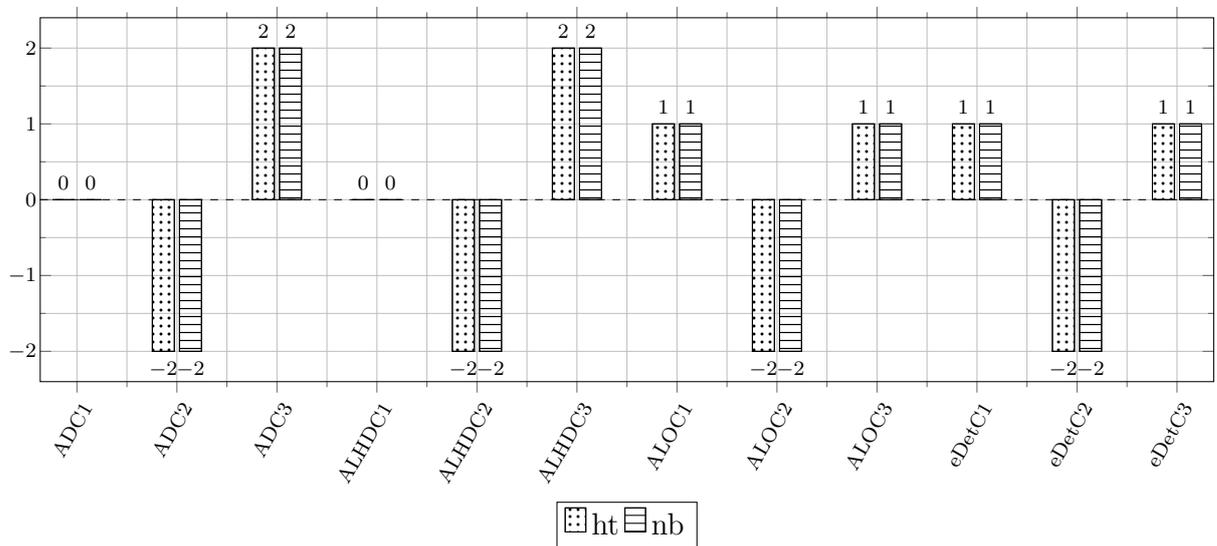
Fonte: O AUTOR (2020).

Os melhores resultados nos conjuntos de dados avaliados foram obtidos pelos comitês ADC₁, ADC₃, ALHDC₁, ALHDC₃, ALOC₁, ALOC₂, ALOC₂ e eDetC₂, ambos com rank dos igual ou acima de zero, mostrados na Figura 38 (b). Apesar disso, vale a pena destacar que o desempenho deles foram bons na maioria das situações com o HT e NB. A Figura 38 (b) mostra a diferença média calculada, levando em consideração as hipóteses testadas com HT e NB. Os comitês ADC₃, ALHDC₃ e eDetC₂ com o HT e NB, forneceram os melhores desempenhos e foram estatisticamente superiores aos demais comitês de mesmo método. Tratando-se do ALO, os comitês ALOC₁, ALOC₂ e ALOC₂ foram iguais estatisticamente com o HT e NB.

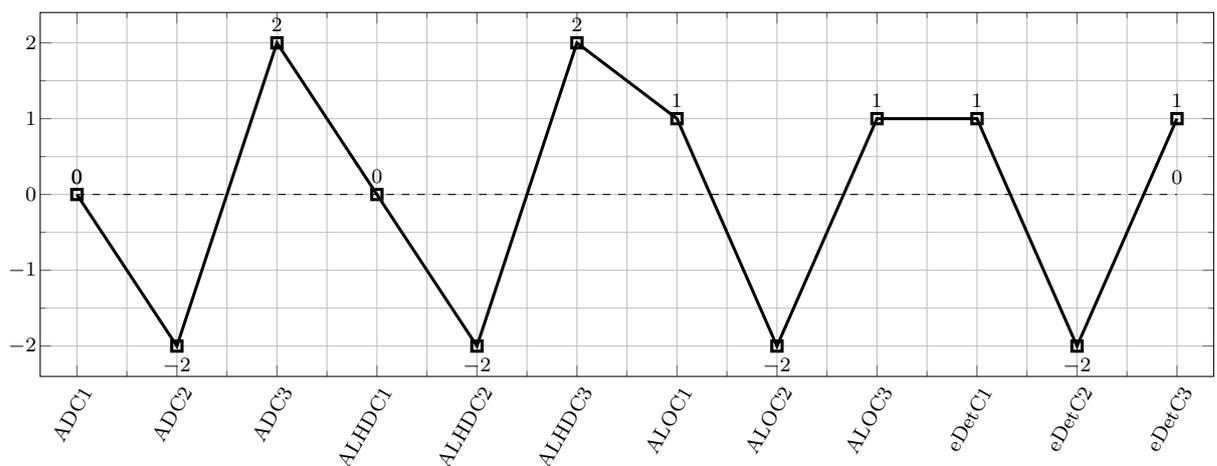
As Tabelas 36 e 37 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês ADC₃, ALHDC₃, ALOC₁, ALOC₂, eDetC₁ e eDetC₂ obtiveram os melhores resultados com o HT e NB. De acordo como os resultados das médias das acurácias e do rank dos , o mapa de calor mostra uma grande e intensa concentração de vermelho nas linhas dos métodos AD e ALHDC₂ com o HT e NB, indicando que estes obtiveram as variações com os piores resultados em relação a seus respectivos métodos.

Figura 39 – Resumo dos ranks dos MCCs com HT e NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



Fonte: O AUTOR (2020).

Previsivelmente, os resultados obtidos com AD e ALHD foram ruins em comparação com o ALO e e-Detector, pois as acurácias também não foram boas. Além disso, os valores médios obtidos com MTR são relativamente grandes e aumentam na mesma

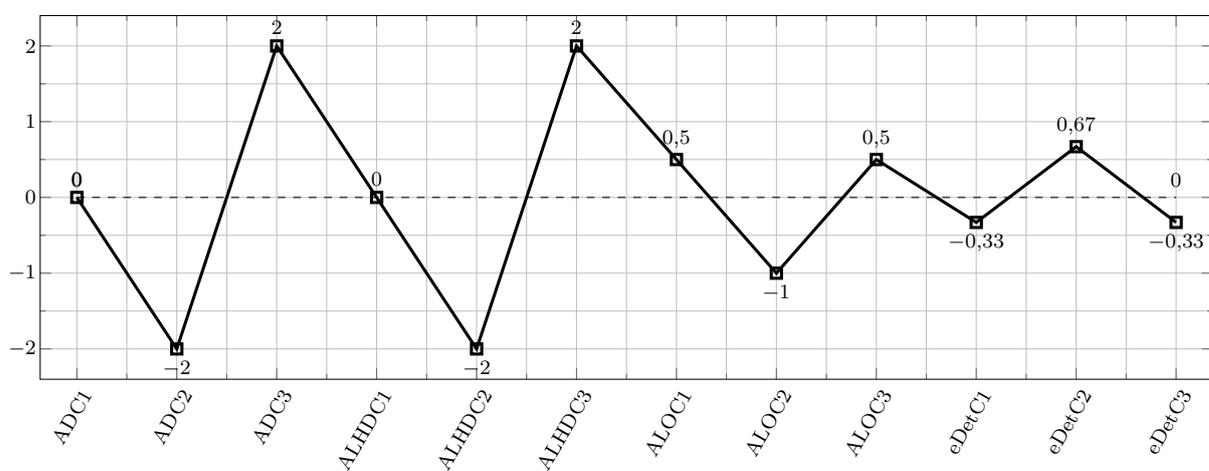
proporção que os conceitos crescem, assim sendo, apesar do ALHD ter conseguido bons resultados com a parametrização ALHDC₃ no acumulado geral os resultados são inferiores aos obtidos pelo *e*-Detector e ALO. Assim, observando em termos estatísticos a Figura 39 (a) mostra o resumo dos resultados com HT e NB apresentados no Apêndice C nas Tabelas 42 e 43. Em termos das hipóteses testadas, os comitês ADC₃, ALHDC₃, ALOC₁, ALOC₂, eDetC₁ e eDetC₃ com o HT e NB, forneceram os melhores desempenhos e foram estatisticamente superiores aos demais comitês de mesmo método. Neste caso, seis dos doze comitês avaliados obtiveram rank com valor positivo e diferentes de zero.

Nota-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês ADC₃, ALHDC₃, ALOC₁, ALOC₂, eDetC₁ e eDetC₃, todos com avaliação acima de zero, mostrados na Figura 39 (b). Assim, é possível observar que os comitês ALOC₁ e ALOC₂ com o ALO, eDetC₁ e eDetC₃ com o *e*-Detector obtiveram médias no desempenho acumulado iguais para cada um dos métodos.

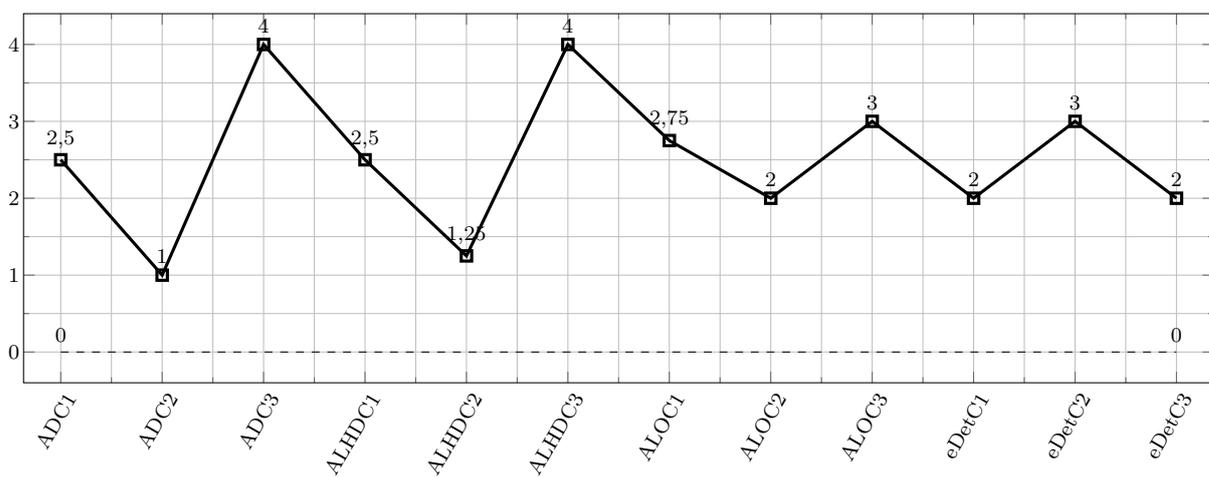
A Figura 40 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 40 (a) é feito o rank da média das diferenças. A Figura 40 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 40 (a) ou no da Figura 40 (b), os comitês de maior valor absoluto foram ADC₃, ALHDC₃, ALOC₁, ALOC₂ e eDetC₂. Vale ressaltar que por meio do rank da média (Figura 40 (a)) é possível visualizar os números negativos que mostram quais comitês obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um comitê menos eficiente nesses cenários.

Figura 40 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando as comparações com os métodos de múltipla análise dos trabalhos relacionados.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.



Fonte: O AUTOR (2020).

Tabela 32 – Médias das acurácias em percentual (%) utilizando HT. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDefC1 | eDefC2 | eDefC3 |
|--------------|-------|-------|-------|--------------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 73.71 | 73.64 | 74.72 | 81.31 | 75.74 | 86.96 | 86.96 | 86.96 | 86.96 | 81.34 | 82.04 | 81.34 |
| A-20k-LED | 63.21 | 63.21 | 63.31 | 64.57 | 63.75 | 68.67 | 68.67 | 68.55 | 68.67 | 70.37 | 70.54 | 70.37 |
| A-20k-Mix | 60.97 | 60.14 | 61.33 | 90.76 | 63.69 | 90.76 | 90.76 | 90.79 | 90.76 | 83.55 | 83.41 | 83.55 |
| A-20k-RBF | 72.34 | 72.34 | 72.45 | 72.26 | 72.73 | 74.84 | 74.84 | 74.78 | 74.84 | 73.09 | 73.12 | 73.09 |
| A-20k-Sine | 62.20 | 62.06 | 62.40 | 89.71 | 65.30 | 89.73 | 89.73 | 89.76 | 89.73 | 79.26 | 79.43 | 79.26 |
| A-20k-Wav | 77.53 | 77.53 | 77.67 | 77.74 | 77.77 | 79.21 | 79.21 | 79.26 | 79.21 | 78.49 | 78.53 | 78.49 |
| A-130k-Agr | 77.47 | 76.29 | 77.81 | 89.82 | 79.00 | 91.47 | 91.47 | 91.27 | 91.47 | 89.98 | 90.04 | 89.98 |
| A-130k-LED | 66.81 | 66.81 | 66.88 | 68.35 | 67.20 | 70.00 | 70.00 | 69.93 | 70.00 | 72.04 | 72.09 | 72.04 |
| A-130k-Mix | 61.26 | 60.73 | 61.38 | 92.58 | 63.63 | 92.58 | 92.58 | 92.57 | 92.58 | 70.47 | 70.84 | 70.47 |
| A-130k-RBF | 77.14 | 77.18 | 77.33 | 77.58 | 77.57 | 78.01 | 78.01 | 77.91 | 78.01 | 78.51 | 78.61 | 78.51 |
| A-130k-Sine | 64.75 | 64.49 | 64.75 | 93.33 | 66.92 | 93.33 | 93.33 | 93.30 | 93.33 | 81.71 | 81.83 | 81.71 |
| A-130k-Wav | 78.99 | 79.01 | 79.08 | 79.19 | 79.14 | 79.70 | 79.70 | 79.70 | 79.70 | 79.30 | 79.31 | 79.30 |
| A-890k-Agr | 77.86 | 77.14 | 78.86 | 95.27 | 80.84 | 95.15 | 95.15 | 94.89 | 95.15 | 94.99 | 94.97 | 94.99 |
| A-890k-LED | 68.12 | 68.12 | 68.44 | 71.66 | 69.92 | 70.35 | 70.35 | 70.36 | 70.35 | 72.39 | 72.44 | 72.39 |
| A-890k-Mix | 61.30 | 60.90 | 61.30 | 94.96 | 66.19 | 94.96 | 94.96 | 94.95 | 94.96 | 74.56 | 74.60 | 74.56 |
| A-890k-RBF | 82.27 | 82.01 | 82.45 | 84.98 | 82.78 | 83.47 | 83.47 | 82.61 | 83.47 | 84.66 | 84.43 | 84.66 |
| A-890k-Sine | 63.51 | 62.97 | 63.84 | 96.97 | 65.68 | 96.97 | 96.97 | 96.93 | 96.97 | 84.91 | 84.83 | 84.91 |
| A-890k-Wav | 81.62 | 81.62 | 81.54 | 81.59 | 81.55 | 80.85 | 80.85 | 80.63 | 80.85 | 81.75 | 81.58 | 81.75 |
| G-20k-Agr | 73.74 | 73.39 | 73.83 | 79.68 | 73.95 | 85.23 | 85.23 | 85.42 | 85.23 | 80.14 | 81.09 | 80.14 |
| G-20k-LED | 63.20 | 63.20 | 63.20 | 64.60 | 63.71 | 68.30 | 68.30 | 68.31 | 68.30 | 69.91 | 69.86 | 69.91 |
| G-20k-Mix | 60.79 | 59.72 | 61.76 | 86.84 | 62.93 | 87.43 | 87.43 | 87.30 | 87.43 | 85.24 | 87.01 | 85.24 |
| G-20k-RBF | 72.11 | 72.11 | 72.14 | 72.07 | 72.14 | 73.83 | 73.83 | 73.82 | 73.83 | 72.88 | 72.95 | 72.88 |
| G-20k-Sine | 62.57 | 61.30 | 63.75 | 78.48 | 65.68 | 86.70 | 86.70 | 86.69 | 86.70 | 84.04 | 85.67 | 84.04 |
| G-20k-Wav | 77.52 | 77.52 | 77.58 | 77.64 | 77.66 | 78.77 | 78.77 | 78.82 | 78.77 | 78.27 | 78.39 | 78.27 |
| G-130k-Agr | 77.19 | 76.22 | 78.81 | 90.05 | 79.18 | 91.19 | 91.19 | 90.99 | 91.19 | 90.50 | 90.74 | 90.50 |
| G-130k-LED | 66.81 | 66.81 | 66.87 | 68.21 | 66.99 | 69.87 | 69.87 | 69.85 | 69.87 | 72.02 | 72.05 | 72.02 |
| G-130k-Mix | 62.00 | 61.27 | 62.09 | 92.03 | 63.94 | 92.04 | 92.04 | 91.99 | 92.04 | 90.93 | 92.06 | 90.93 |
| G-130k-RBF | 77.25 | 77.15 | 77.33 | 77.48 | 77.44 | 77.79 | 77.79 | 77.60 | 77.79 | 78.64 | 78.81 | 78.64 |
| G-130k-Sine | 66.07 | 64.44 | 66.51 | 92.62 | 68.01 | 92.78 | 92.78 | 92.72 | 92.78 | 92.23 | 92.74 | 92.23 |
| G-130k-Wav | 78.99 | 78.99 | 79.02 | 79.10 | 79.00 | 79.71 | 79.71 | 79.60 | 79.71 | 79.36 | 79.33 | 79.36 |
| G-890k-Agr | 77.98 | 77.14 | 79.01 | 95.23 | 80.87 | 95.11 | 95.11 | 94.84 | 95.11 | 95.11 | 95.13 | 95.11 |
| G-890k-LED | 68.12 | 68.12 | 68.70 | 71.78 | 69.92 | 70.34 | 70.34 | 70.34 | 70.34 | 72.37 | 72.44 | 72.37 |
| G-890k-Mix | 62.68 | 62.01 | 62.89 | 94.85 | 66.73 | 94.85 | 94.85 | 94.84 | 94.85 | 94.88 | 94.88 | 94.88 |
| G-890k-RBF | 82.58 | 82.00 | 82.73 | 84.94 | 82.93 | 83.07 | 83.07 | 82.43 | 83.07 | 84.76 | 84.49 | 84.76 |
| G-890k-Sine | 65.77 | 64.28 | 66.77 | 96.85 | 68.92 | 96.86 | 96.86 | 96.80 | 96.86 | 96.90 | 96.86 | 96.90 |
| G-890k-Wav | 81.61 | 81.62 | 81.52 | 81.61 | 81.53 | 80.86 | 80.86 | 80.54 | 80.86 | 81.77 | 81.56 | 81.77 |
| Airlines | 65.44 | 65.44 | 65.44 | 65.86 | 66.08 | 65.67 | 65.67 | 65.62 | 65.67 | 66.45 | 66.46 | 66.45 |
| Connect4 | 73.66 | 73.66 | 74.00 | 73.77 | 73.36 | 75.10 | 75.10 | 75.05 | 75.10 | 74.74 | 74.91 | 74.74 |
| Outdoor | 57.98 | 56.99 | 58.42 | 57.57 | 57.98 | 61.14 | 61.14 | 60.07 | 61.14 | 60.67 | 61.27 | 60.67 |
| Rialto | 31.10 | 31.10 | 31.10 | 29.96 | 28.61 | 39.98 | 39.98 | 41.31 | 39.98 | 35.11 | 38.81 | 35.11 |
| SpamData | 91.57 | 91.57 | 91.57 | 92.02 | 91.57 | 92.16 | 92.16 | 92.89 | 92.16 | 92.04 | 92.32 | 92.04 |
| WineRed | 50.96 | 50.96 | 50.96 | 50.96 | 50.96 | 52.24 | 52.24 | 52.24 | 52.24 | 51.33 | 51.33 | 51.33 |
| WineWhite | 45.13 | 45.13 | 45.13 | 45.13 | 45.13 | 47.14 | 47.14 | 47.13 | 47.14 | 44.80 | 44.80 | 44.80 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 2 | 0 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 1 | 0 | 1 |
| Média Postos | 2,13 | 2,62 | 1,26 | 1,91 | 2,77 | 1,33 | 1,74 | 2,51 | 1,74 | 2,24 | 1,51 | 2,24 |
| Posição | 2 | 3 | 1 | 2 | 3 | 1 | 1,5 | 3 | 1,5 | 2,5 | 1 | 2,5 |

Fonte: O AUTOR (2020).

Tabela 33 – Médias das acurácias em percentual (%) utilizando NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------|-------|-------|-------|--------------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 67.85 | 67.85 | 68.42 | 77.13 | 70.19 | 86.99 | 86.99 | 87.03 | 86.99 | 79.54 | 79.51 | 79.54 |
| A-20k-LED | 60.01 | 60.01 | 60.11 | 62.03 | 60.96 | 69.74 | 69.74 | 69.74 | 69.74 | 68.97 | 69.33 | 68.97 |
| A-20k-Mix | 57.99 | 57.99 | 57.99 | 91.13 | 58.09 | 91.13 | 91.13 | 91.18 | 91.13 | 76.24 | 75.18 | 76.24 |
| A-20k-RBF | 63.00 | 63.00 | 63.28 | 64.77 | 63.80 | 74.86 | 74.86 | 74.83 | 74.86 | 70.57 | 70.55 | 70.57 |
| A-20k-Sine | 59.01 | 58.44 | 59.24 | 83.16 | 59.74 | 89.15 | 89.15 | 89.18 | 89.15 | 69.85 | 70.47 | 69.85 |
| A-20k-Wav | 75.37 | 75.44 | 75.73 | 75.78 | 75.83 | 79.70 | 79.70 | 79.71 | 79.70 | 77.82 | 78.23 | 77.82 |
| A-130k-Agr | 68.80 | 68.78 | 69.12 | 84.55 | 70.73 | 88.72 | 88.72 | 88.71 | 88.72 | 86.93 | 86.90 | 86.93 |
| A-130k-LED | 59.83 | 59.83 | 60.33 | 66.63 | 62.16 | 71.94 | 71.94 | 71.90 | 71.94 | 73.08 | 73.03 | 73.08 |
| A-130k-Mix | 58.34 | 58.34 | 58.34 | 91.93 | 60.03 | 91.93 | 91.93 | 91.94 | 91.93 | 66.31 | 65.30 | 66.31 |
| A-130k-RBF | 62.67 | 62.67 | 63.61 | 66.33 | 63.98 | 75.25 | 75.25 | 75.25 | 75.25 | 74.38 | 74.36 | 74.38 |
| A-130k-Sine | 59.88 | 58.82 | 59.88 | 83.88 | 59.88 | 89.48 | 89.48 | 89.48 | 89.48 | 62.70 | 62.54 | 62.70 |
| A-130k-Wav | 75.20 | 75.20 | 75.48 | 76.03 | 76.00 | 80.32 | 80.32 | 80.31 | 80.32 | 79.68 | 79.77 | 79.68 |
| A-890k-Agr | 69.00 | 69.00 | 69.34 | 86.07 | 71.50 | 89.10 | 89.10 | 89.09 | 89.10 | 88.99 | 88.99 | 88.99 |
| A-890k-LED | 59.89 | 59.89 | 60.61 | 68.54 | 63.85 | 72.49 | 72.49 | 72.44 | 72.49 | 73.79 | 73.78 | 73.79 |
| A-890k-Mix | 58.41 | 58.41 | 58.41 | 92.07 | 58.41 | 92.07 | 92.07 | 92.07 | 92.07 | 59.02 | 60.27 | 59.02 |
| A-890k-RBF | 62.65 | 62.65 | 63.20 | 68.21 | 64.93 | 75.36 | 75.36 | 75.35 | 75.36 | 75.33 | 75.32 | 75.33 |
| A-890k-Sine | 59.20 | 58.85 | 59.47 | 83.37 | 59.47 | 89.50 | 89.50 | 89.50 | 89.50 | 62.74 | 75.73 | 62.74 |
| A-890k-Wav | 75.19 | 75.19 | 76.09 | 78.87 | 76.94 | 80.37 | 80.37 | 80.37 | 80.37 | 80.29 | 80.30 | 80.29 |
| G-20k-Agr | 67.81 | 67.81 | 67.93 | 73.43 | 68.38 | 84.23 | 84.23 | 84.72 | 84.23 | 77.43 | 78.04 | 77.43 |
| G-20k-LED | 60.00 | 60.00 | 60.09 | 61.06 | 60.23 | 69.26 | 69.26 | 69.35 | 69.26 | 68.83 | 69.00 | 68.83 |
| G-20k-Mix | 58.26 | 58.20 | 58.29 | 86.96 | 58.70 | 87.83 | 87.83 | 87.69 | 87.83 | 86.39 | 87.63 | 86.39 |
| G-20k-RBF | 63.02 | 63.02 | 63.46 | 64.80 | 63.63 | 73.78 | 73.78 | 73.83 | 73.78 | 70.92 | 71.21 | 70.92 |
| G-20k-Sine | 58.69 | 58.69 | 59.46 | 75.98 | 59.83 | 86.75 | 86.75 | 86.68 | 86.75 | 83.87 | 85.66 | 83.87 |
| G-20k-Wav | 75.34 | 75.34 | 75.43 | 75.50 | 75.56 | 78.94 | 78.94 | 79.00 | 78.94 | 77.46 | 77.78 | 77.46 |
| G-130k-Agr | 68.80 | 68.78 | 69.12 | 84.43 | 70.05 | 88.33 | 88.33 | 88.32 | 88.33 | 87.03 | 87.08 | 87.03 |
| G-130k-LED | 59.83 | 59.83 | 60.58 | 66.30 | 61.48 | 71.87 | 71.87 | 71.81 | 71.87 | 73.08 | 73.11 | 73.08 |
| G-130k-Mix | 58.36 | 58.36 | 58.36 | 91.39 | 58.38 | 91.40 | 91.40 | 91.39 | 91.40 | 89.39 | 91.40 | 89.39 |
| G-130k-RBF | 62.75 | 62.67 | 63.68 | 66.09 | 63.80 | 75.05 | 75.05 | 75.06 | 75.05 | 74.65 | 74.70 | 74.65 |
| G-130k-Sine | 59.04 | 58.83 | 59.04 | 79.94 | 59.00 | 89.19 | 89.19 | 89.19 | 89.19 | 86.92 | 89.10 | 86.92 |
| G-130k-Wav | 75.20 | 75.20 | 75.79 | 76.16 | 76.07 | 80.10 | 80.10 | 80.12 | 80.10 | 79.63 | 79.76 | 79.63 |
| G-890k-Agr | 69.00 | 69.00 | 69.30 | 86.25 | 72.50 | 89.02 | 89.02 | 89.00 | 89.02 | 89.04 | 89.05 | 89.04 |
| G-890k-LED | 60.06 | 59.89 | 60.94 | 69.08 | 63.03 | 72.48 | 72.48 | 72.43 | 72.48 | 73.78 | 73.78 | 73.78 |
| G-890k-Mix | 58.41 | 58.41 | 58.41 | 91.99 | 58.42 | 91.99 | 91.99 | 91.99 | 91.99 | 92.00 | 92.00 | 92.00 |
| G-890k-RBF | 62.65 | 62.65 | 64.10 | 68.83 | 66.07 | 75.33 | 75.33 | 75.33 | 75.33 | 75.36 | 75.36 | 75.36 |
| G-890k-Sine | 59.53 | 58.85 | 59.53 | 83.20 | 59.51 | 89.47 | 89.47 | 89.47 | 89.47 | 89.42 | 89.46 | 89.42 |
| G-890k-Wav | 75.19 | 75.19 | 76.22 | 78.78 | 76.91 | 80.34 | 80.34 | 80.33 | 80.34 | 80.30 | 80.32 | 80.30 |
| Airlines | 64.56 | 64.56 | 65.41 | 66.70 | 65.41 | 65.66 | 65.66 | 65.66 | 65.66 | 66.80 | 66.81 | 66.80 |
| Connect4 | 68.55 | 69.46 | 70.88 | 72.85 | 71.63 | 75.07 | 75.07 | 75.05 | 75.07 | 74.41 | 74.47 | 74.41 |
| Outdoor | 58.03 | 57.05 | 58.53 | 57.76 | 58.14 | 61.59 | 61.59 | 60.55 | 61.59 | 60.14 | 61.23 | 60.14 |
| Rialto | 21.59 | 21.59 | 21.59 | 24.46 | 21.59 | 36.84 | 36.84 | 41.31 | 36.84 | 27.40 | 35.86 | 27.40 |
| SpamData | 90.62 | 90.62 | 90.62 | 91.35 | 90.62 | 91.41 | 91.41 | 91.94 | 91.41 | 90.97 | 91.36 | 90.97 |
| WineRed | 47.70 | 47.70 | 47.70 | 47.70 | 47.70 | 47.82 | 47.82 | 47.82 | 47.82 | 47.76 | 47.76 | 47.76 |
| WineWhite | 42.82 | 42.82 | 42.82 | 42.93 | 42.82 | 46.67 | 46.67 | 46.67 | 46.67 | 45.95 | 46.72 | 45.95 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Média Postos | 2,26 | 2,50 | 1,24 | 2,02 | 2,92 | 1,06 | 1,87 | 2,26 | 1,87 | 2,21 | 1,58 | 2,21 |
| Posição | 2 | 3 | 1 | 2 | 3 | 1 | 1,5 | 3 | 1,5 | 2,5 | 1 | 2,5 |

Fonte: O AUTOR (2020).

Tabela 34 – Médias dos MTR utilizando HT. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------|------|------|------|----------------|---------|----------------|----------------|-----------------|----------------|--------------|----------------|--------------|
| A-20k-Agr | 0.01 | 0.00 | 0.11 | 4.42 | 0.26 | 123.84 | 123.84 | 128.19 | 123.84 | 93.69 | 104.02 | 93.69 |
| A-20k-LED | 0.00 | 0.00 | 0.01 | 0.20 | 0.06 | 4.71 | 4.71 | 10.22 | 4.71 | 4.24 | 4.38 | 4.24 |
| A-20k-Mix | 0.03 | 0.00 | 0.05 | 338.74 | 0.17 | 334.99 | 334.99 | 445.32 | 334.99 | 198.63 | 235.04 | 198.63 |
| A-20k-RBF | 0.00 | 0.00 | 0.06 | 0.29 | 0.17 | 31.75 | 31.75 | 35.90 | 31.75 | 30.88 | 35.25 | 30.88 |
| A-20k-Sine | 0.01 | 0.00 | 0.02 | 282.09 | 0.22 | 305.81 | 305.81 | 414.09 | 305.81 | 212.73 | 281.37 | 212.73 |
| A-20k-Wav | 0.00 | 0.01 | 0.03 | 0.10 | 0.09 | 1.77 | 1.77 | 2.14 | 1.77 | 0.99 | 1.18 | 0.99 |
| A-130k-Agr | 0.07 | 0.00 | 0.10 | 95.62 | 0.20 | 965.99 | 965.99 | 1080.12 | 965.99 | 366.78 | 413.64 | 366.78 |
| A-130k-LED | 0.00 | 0.00 | 0.02 | 0.37 | 0.10 | 9.68 | 9.68 | 10.79 | 9.68 | 8.92 | 10.07 | 8.92 |
| A-130k-Mix | 0.06 | 0.00 | 0.08 | 1920.33 | 114.80 | 1920.33 | 1920.33 | 2987.66 | 1920.33 | 1129.24 | 1746.89 | 1129.24 |
| A-130k-RBF | 0.02 | 0.00 | 0.09 | 1.14 | 0.16 | 225.27 | 225.27 | 261.59 | 225.27 | 219.24 | 220.01 | 219.24 |
| A-130k-Sine | 0.02 | 0.01 | 0.02 | 1948.01 | 0.12 | 2022.41 | 2022.41 | 3119.61 | 2022.41 | 960.66 | 1352.86 | 960.66 |
| A-130k-Wav | 0.00 | 0.01 | 0.07 | 0.22 | 0.20 | 43.54 | 43.54 | 30.84 | 43.54 | 88.43 | 64.85 | 88.43 |
| A-890k-Agr | 0.06 | 0.01 | 0.15 | 8775.98 | 0.51 | 7759.54 | 7759.54 | 7591.01 | 7759.54 | 2523.80 | 2330.17 | 2523.80 |
| A-890k-LED | 0.00 | 0.00 | 0.04 | 16.33 | 0.31 | 9.94 | 9.94 | 10.23 | 9.94 | 9.44 | 18.88 | 9.44 |
| A-890k-Mix | 0.03 | 0.01 | 0.03 | 16048.82 | 1587.13 | 15870.89 | 15870.89 | 20681.77 | 15870.89 | 8163.69 | 9724.68 | 8163.69 |
| A-890k-RBF | 0.08 | 0.01 | 0.12 | 791.38 | 0.28 | 1400.76 | 1400.76 | 1187.70 | 1400.76 | 702.05 | 754.97 | 702.05 |
| A-890k-Sine | 0.03 | 0.00 | 0.05 | 16288.62 | 0.19 | 16288.62 | 16288.62 | 21277.01 | 16288.62 | 8918.58 | 11161.84 | 8918.58 |
| A-890k-Wav | 0.00 | 0.00 | 0.18 | 0.53 | 0.25 | 47.74 | 47.74 | 44.11 | 47.74 | 128.59 | 260.30 | 128.59 |
| G-20k-Agr | 0.03 | 0.00 | 0.04 | 1.37 | 0.06 | 9.79 | 9.79 | 5.94 | 9.79 | 7.72 | 7.18 | 7.72 |
| G-20k-LED | 0.00 | 0.00 | 0.00 | 0.14 | 0.02 | 1.64 | 1.64 | 1.56 | 1.64 | 1.60 | 1.61 | 1.60 |
| G-20k-Mix | 0.03 | 0.00 | 0.06 | 3.84 | 0.10 | 6.24 | 6.24 | 6.77 | 6.24 | 12.27 | 15.60 | 12.27 |
| G-20k-RBF | 0.00 | 0.00 | 0.02 | 0.14 | 0.06 | 6.78 | 6.78 | 6.41 | 6.78 | 6.63 | 7.41 | 6.63 |
| G-20k-Sine | 0.05 | 0.00 | 0.11 | 1.27 | 0.20 | 3.24 | 3.24 | 2.00 | 3.24 | 14.46 | 14.42 | 14.46 |
| G-20k-Wav | 0.00 | 0.00 | 0.02 | 0.06 | 0.04 | 0.65 | 0.65 | 0.86 | 0.65 | 0.54 | 0.68 | 0.54 |
| G-130k-Agr | 0.07 | 0.00 | 0.21 | 36.31 | 0.26 | 89.49 | 89.49 | 58.55 | 89.49 | 63.21 | 80.31 | 63.21 |
| G-130k-LED | 0.00 | 0.00 | 0.01 | 0.50 | 0.05 | 2.47 | 2.47 | 3.17 | 2.47 | 3.61 | 3.84 | 3.61 |
| G-130k-Mix | 0.03 | 0.00 | 0.04 | 24.87 | 0.11 | 28.06 | 28.06 | 45.33 | 28.06 | 92.13 | 163.89 | 92.13 |
| G-130k-RBF | 0.03 | 0.00 | 0.08 | 1.30 | 0.16 | 48.00 | 48.00 | 42.75 | 48.00 | 40.13 | 36.88 | 40.13 |
| G-130k-Sine | 0.06 | 0.00 | 0.07 | 3.75 | 0.16 | 21.29 | 21.29 | 13.45 | 21.29 | 122.01 | 128.37 | 122.01 |
| G-130k-Wav | 0.00 | 0.00 | 0.03 | 0.13 | 0.07 | 3.35 | 3.35 | 5.43 | 3.35 | 7.08 | 12.54 | 7.08 |
| G-890k-Agr | 0.06 | 0.01 | 0.13 | 840.08 | 0.36 | 799.73 | 799.73 | 364.31 | 799.73 | 536.85 | 434.47 | 536.85 |
| G-890k-LED | 0.00 | 0.00 | 0.07 | 1.17 | 0.32 | 3.13 | 3.13 | 2.44 | 3.13 | 3.11 | 6.61 | 3.11 |
| G-890k-Mix | 0.02 | 0.00 | 0.03 | 232.61 | 0.18 | 383.43 | 383.43 | 365.06 | 383.43 | 1593.44 | 1766.60 | 1593.44 |
| G-890k-RBF | 0.17 | 0.01 | 0.27 | 178.92 | 0.34 | 246.21 | 246.21 | 94.75 | 246.21 | 213.02 | 271.34 | 213.02 |
| G-890k-Sine | 0.05 | 0.00 | 0.09 | 34.17 | 0.24 | 46.50 | 46.50 | 49.04 | 46.50 | 1272.62 | 1316.23 | 1272.62 |
| G-890k-Wav | 0.01 | 0.00 | 0.22 | 0.60 | 0.34 | 17.03 | 17.03 | 8.17 | 17.03 | 74.96 | 75.20 | 74.96 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Média Postos | 2,17 | 2,78 | 1,06 | 1,83 | 3,00 | 1,17 | 2,06 | 1,89 | 2,06 | 2,33 | 1,33 | 2,33 |
| Posição | 2 | 3 | 1 | 2 | 3 | 1 | 2,5 | 1 | 2,5 | 2,5 | 1 | 2,5 |

Fonte: O AUTOR (2020).

Tabela 35 – Médias dos MTR utilizando NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------|------|------|------|----------|--------|-----------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|
| A-20k-Agr | 0.00 | 0.00 | 0.03 | 2.57 | 2.74 | 126.37 | 126.37 | 144.49 | 126.37 | 97.77 | 114.84 | 97.77 |
| A-20k-LED | 0.00 | 0.00 | 0.01 | 0.12 | 0.07 | 30.52 | 30.52 | 32.06 | 30.52 | 26.21 | 28.85 | 26.21 |
| A-20k-Mix | 0.00 | 0.00 | 0.00 | 341.21 | 0.02 | 341.21 | 341.21 | 537.58 | 341.21 | 299.60 | 487.89 | 299.60 |
| A-20k-RBF | 0.00 | 0.00 | 0.02 | 0.23 | 0.07 | 24.68 | 24.68 | 28.88 | 24.68 | 18.33 | 26.86 | 18.33 |
| A-20k-Sine | 0.05 | 0.00 | 0.07 | 157.09 | 0.10 | 311.75 | 311.75 | 381.05 | 311.75 | 253.89 | 321.49 | 253.89 |
| A-20k-Wav | 0.00 | 0.01 | 0.06 | 0.13 | 0.08 | 44.00 | 44.00 | 60.10 | 44.00 | 2.96 | 4.43 | 2.96 |
| A-130k-Agr | 0.01 | 0.00 | 0.02 | 22.23 | 0.13 | 656.66 | 656.66 | 651.17 | 656.66 | 333.36 | 336.73 | 333.36 |
| A-130k-LED | 0.00 | 0.00 | 0.03 | 0.41 | 0.13 | 25.69 | 25.69 | 27.86 | 25.69 | 60.82 | 64.09 | 60.82 |
| A-130k-Mix | 0.00 | 0.00 | 0.00 | 2234.19 | 96.40 | 2234.19 | 2234.19 | 2517.51 | 2234.19 | 1775.55 | 2347.58 | 1775.55 |
| A-130k-RBF | 0.00 | 0.00 | 0.09 | 0.48 | 0.13 | 87.01 | 87.01 | 111.43 | 87.01 | 76.55 | 92.68 | 76.55 |
| A-130k-Sine | 0.09 | 0.00 | 0.09 | 1101.99 | 0.09 | 2112.52 | 2112.52 | 2307.84 | 2112.52 | 1646.80 | 1975.02 | 1646.80 |
| A-130k-Wav | 0.00 | 0.00 | 0.03 | 0.18 | 0.12 | 413.18 | 413.18 | 415.80 | 413.18 | 90.27 | 107.76 | 90.27 |
| A-890k-Agr | 0.00 | 0.00 | 0.02 | 28.10 | 0.69 | 4096.23 | 4096.23 | 3841.47 | 4096.23 | 1671.58 | 1495.59 | 1671.58 |
| A-890k-LED | 0.00 | 0.00 | 0.05 | 0.91 | 0.34 | 34.23 | 34.23 | 34.04 | 34.23 | 61.26 | 66.94 | 61.26 |
| A-890k-Mix | 0.00 | 0.00 | 0.00 | 15050.81 | 0.00 | 15050.81 | 15050.81 | 16064.36 | 15050.81 | 12542.27 | 15795.99 | 12542.27 |
| A-890k-RBF | 0.00 | 0.00 | 0.03 | 1.15 | 0.31 | 226.77 | 226.77 | 259.22 | 226.77 | 232.65 | 276.83 | 232.65 |
| A-890k-Sine | 0.04 | 0.00 | 0.06 | 6385.81 | 0.06 | 13016.89 | 13016.89 | 12621.18 | 13016.89 | 9644.38 | 6401.19 | 9644.38 |
| A-890k-Wav | 0.00 | 0.00 | 0.09 | 27.54 | 0.21 | 1028.78 | 1028.78 | 774.54 | 1028.78 | 1221.40 | 1137.73 | 1221.40 |
| G-20k-Agr | 0.00 | 0.00 | 0.01 | 0.90 | 0.04 | 6.39 | 6.39 | 9.53 | 6.39 | 9.81 | 7.78 | 9.81 |
| G-20k-LED | 0.00 | 0.00 | 0.01 | 0.07 | 0.02 | 3.05 | 3.05 | 3.90 | 3.05 | 6.12 | 8.43 | 6.12 |
| G-20k-Mix | 0.01 | 0.00 | 0.01 | 3.35 | 0.03 | 3.42 | 3.42 | 5.82 | 3.42 | 16.09 | 19.60 | 16.09 |
| G-20k-RBF | 0.00 | 0.00 | 0.03 | 0.19 | 0.05 | 4.81 | 4.81 | 5.86 | 4.81 | 6.34 | 7.16 | 6.34 |
| G-20k-Sine | 0.00 | 0.00 | 0.06 | 1.18 | 0.08 | 3.24 | 3.24 | 1.99 | 3.24 | 10.87 | 13.97 | 10.87 |
| G-20k-Wav | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.43 | 0.43 | 0.51 | 0.43 | 0.52 | 0.64 | 0.52 |
| G-130k-Agr | 0.01 | 0.00 | 0.02 | 3.24 | 0.07 | 45.44 | 45.44 | 35.86 | 45.44 | 47.94 | 40.59 | 47.94 |
| G-130k-LED | 0.00 | 0.00 | 0.05 | 0.42 | 0.10 | 6.36 | 6.36 | 7.26 | 6.36 | 15.45 | 19.91 | 15.45 |
| G-130k-Mix | 0.00 | 0.00 | 0.00 | 19.94 | 0.03 | 21.78 | 21.78 | 44.60 | 21.78 | 99.09 | 150.85 | 99.09 |
| G-130k-RBF | 0.01 | 0.00 | 0.08 | 0.44 | 0.09 | 17.75 | 17.75 | 21.74 | 17.75 | 25.22 | 25.37 | 25.22 |
| G-130k-Sine | 0.03 | 0.00 | 0.03 | 1.48 | 0.05 | 13.94 | 13.94 | 22.55 | 13.94 | 110.73 | 135.54 | 110.73 |
| G-130k-Wav | 0.00 | 0.00 | 0.09 | 0.17 | 0.11 | 72.67 | 72.67 | 62.99 | 72.67 | 13.13 | 23.76 | 13.13 |
| G-890k-Agr | 0.00 | 0.00 | 0.02 | 51.69 | 0.46 | 220.29 | 220.29 | 295.94 | 220.29 | 280.65 | 280.37 | 280.65 |
| G-890k-LED | 0.01 | 0.00 | 0.07 | 4.81 | 0.24 | 9.85 | 9.85 | 9.69 | 9.85 | 17.56 | 24.02 | 17.56 |
| G-890k-Mix | 0.00 | 0.00 | 0.00 | 210.57 | 0.05 | 268.25 | 268.25 | 278.39 | 268.25 | 1451.34 | 1527.70 | 1451.34 |
| G-890k-RBF | 0.00 | 0.00 | 0.11 | 1.33 | 0.37 | 19.60 | 19.60 | 17.82 | 19.60 | 65.13 | 68.65 | 65.13 |
| G-890k-Sine | 0.08 | 0.00 | 0.08 | 19.68 | 0.09 | 38.22 | 38.22 | 37.29 | 38.22 | 1044.14 | 1060.42 | 1044.14 |
| G-890k-Wav | 0.00 | 0.00 | 0.11 | 3.19 | 0.19 | 83.70 | 83.70 | 32.01 | 83.70 | 156.12 | 150.44 | 156.12 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Média Postos | 2,25 | 2,56 | 1,19 | 1,99 | 2,97 | 1,04 | 2,17 | 1,67 | 2,17 | 2,31 | 1,39 | 2,31 |
| Posição | 2 | 3 | 1 | 2 | 3 | 1 | 2,5 | 1 | 2,5 | 2,5 | 1 | 2,5 |

Fonte: O AUTOR (2020).

Tabela 36 – Médias calculadas do MCC utilizando HT. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------|------|------|------|-------------|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| A-20k-Agr | 0.01 | 0.00 | 0.13 | 0.57 | 0.26 | 0.87 | 0.87 | 0.86 | 0.87 | 0.74 | 0.73 | 0.74 |
| A-20k-LED | 0.00 | 0.00 | 0.01 | 0.02 | 0.04 | 0.19 | 0.19 | 0.20 | 0.19 | 0.18 | 0.18 | 0.18 |
| A-20k-Mix | 0.05 | 0.00 | 0.07 | 0.95 | 0.18 | 0.95 | 0.95 | 0.95 | 0.95 | 0.74 | 0.73 | 0.74 |
| A-20k-RBF | 0.00 | 0.00 | 0.06 | 0.08 | 0.17 | 0.62 | 0.62 | 0.60 | 0.62 | 0.59 | 0.57 | 0.59 |
| A-20k-Sine | 0.01 | 0.00 | 0.03 | 0.97 | 0.17 | 0.97 | 0.97 | 0.96 | 0.97 | 0.79 | 0.78 | 0.79 |
| A-20k-Wav | 0.00 | 0.01 | 0.04 | 0.05 | 0.09 | 0.28 | 0.28 | 0.28 | 0.28 | 0.40 | 0.40 | 0.40 |
| A-130k-Agr | 0.10 | 0.00 | 0.13 | 0.71 | 0.24 | 0.82 | 0.82 | 0.80 | 0.82 | 0.70 | 0.69 | 0.70 |
| A-130k-LED | 0.00 | 0.00 | 0.00 | 0.01 | 0.03 | 0.08 | 0.08 | 0.09 | 0.08 | 0.09 | 0.09 | 0.09 |
| A-130k-Mix | 0.07 | 0.00 | 0.09 | 0.92 | 0.21 | 0.92 | 0.92 | 0.92 | 0.92 | 0.71 | 0.70 | 0.71 |
| A-130k-RBF | 0.03 | 0.00 | 0.09 | 0.26 | 0.19 | 0.46 | 0.46 | 0.45 | 0.46 | 0.48 | 0.49 | 0.48 |
| A-130k-Sine | 0.03 | 0.01 | 0.03 | 0.96 | 0.17 | 0.96 | 0.96 | 0.94 | 0.96 | 0.72 | 0.71 | 0.72 |
| A-130k-Wav | 0.00 | 0.01 | 0.08 | 0.07 | 0.11 | 0.24 | 0.24 | 0.21 | 0.24 | 0.35 | 0.27 | 0.35 |
| A-890k-Agr | 0.09 | 0.01 | 0.12 | 0.95 | 0.17 | 0.83 | 0.83 | 0.73 | 0.83 | 0.71 | 0.68 | 0.71 |
| A-890k-LED | 0.00 | 0.00 | 0.00 | 0.03 | 0.02 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 |
| A-890k-Mix | 0.05 | 0.01 | 0.05 | 0.93 | 0.26 | 0.93 | 0.93 | 0.92 | 0.93 | 0.70 | 0.70 | 0.70 |
| A-890k-RBF | 0.11 | 0.01 | 0.13 | 0.63 | 0.19 | 0.33 | 0.33 | 0.28 | 0.33 | 0.42 | 0.38 | 0.42 |
| A-890k-Sine | 0.04 | 0.00 | 0.06 | 0.97 | 0.14 | 0.97 | 0.97 | 0.95 | 0.97 | 0.75 | 0.74 | 0.75 |
| A-890k-Wav | 0.00 | 0.00 | 0.04 | 0.05 | 0.06 | 0.13 | 0.13 | 0.11 | 0.13 | 0.20 | 0.22 | 0.20 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 0 | 2 | 0 |
| Média Postos | 2,08 | 2,64 | 1,28 | 2,00 | 2,67 | 1,33 | 1,72 | 2,56 | 1,72 | 1,72 | 2,56 | 1,72 |
| Posição | 8 | 11 | 1 | 7 | 12 | 2 | 4,5 | 9,5 | 4,5 | 4,5 | 9,5 | 4,5 |

Fonte: O AUTOR (2020).

Tabela 37 – Médias calculadas do MCC utilizando NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------|------|------|------|-------------|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| A-20k-Agr | 0.00 | 0.00 | 0.05 | 0.47 | 0.14 | 0.87 | 0.87 | 0.87 | 0.87 | 0.76 | 0.74 | 0.76 |
| A-20k-LED | 0.00 | 0.00 | 0.00 | 0.02 | 0.01 | 0.32 | 0.32 | 0.33 | 0.32 | 0.40 | 0.38 | 0.40 |
| A-20k-Mix | 0.00 | 0.00 | 0.00 | 0.98 | 0.03 | 0.98 | 0.98 | 0.98 | 0.98 | 0.88 | 0.87 | 0.88 |
| A-20k-RBF | 0.00 | 0.00 | 0.04 | 0.07 | 0.11 | 0.65 | 0.65 | 0.61 | 0.65 | 0.56 | 0.54 | 0.56 |
| A-20k-Sine | 0.06 | 0.00 | 0.06 | 0.85 | 0.09 | 0.98 | 0.98 | 0.98 | 0.98 | 0.81 | 0.80 | 0.81 |
| A-20k-Wav | 0.00 | 0.01 | 0.07 | 0.07 | 0.09 | 0.67 | 0.67 | 0.72 | 0.67 | 0.50 | 0.53 | 0.50 |
| A-130k-Agr | 0.00 | 0.00 | 0.01 | 0.70 | 0.09 | 0.61 | 0.61 | 0.59 | 0.61 | 0.56 | 0.56 | 0.56 |
| A-130k-LED | 0.00 | 0.00 | 0.00 | 0.05 | 0.02 | 0.14 | 0.14 | 0.15 | 0.14 | 0.23 | 0.23 | 0.23 |
| A-130k-Mix | 0.00 | 0.00 | 0.00 | 0.98 | 0.06 | 0.98 | 0.98 | 0.98 | 0.98 | 0.79 | 0.79 | 0.79 |
| A-130k-RBF | 0.00 | 0.00 | 0.09 | 0.14 | 0.12 | 0.33 | 0.33 | 0.33 | 0.33 | 0.33 | 0.32 | 0.33 |
| A-130k-Sine | 0.08 | 0.00 | 0.08 | 0.86 | 0.08 | 0.97 | 0.97 | 0.96 | 0.97 | 0.77 | 0.76 | 0.77 |
| A-130k-Wav | 0.00 | 0.00 | 0.04 | 0.04 | 0.07 | 0.65 | 0.65 | 0.63 | 0.65 | 0.48 | 0.46 | 0.48 |
| A-890k-Agr | 0.00 | 0.00 | 0.01 | 0.70 | 0.05 | 0.28 | 0.28 | 0.27 | 0.28 | 0.29 | 0.28 | 0.29 |
| A-890k-LED | 0.00 | 0.00 | 0.00 | 0.02 | 0.01 | 0.07 | 0.07 | 0.06 | 0.07 | 0.08 | 0.08 | 0.08 |
| A-890k-Mix | 0.00 | 0.00 | 0.00 | 0.98 | 0.00 | 0.98 | 0.98 | 0.97 | 0.98 | 0.77 | 0.76 | 0.77 |
| A-890k-RBF | 0.00 | 0.00 | 0.03 | 0.05 | 0.05 | 0.14 |
| A-890k-Sine | 0.04 | 0.00 | 0.04 | 0.79 | 0.04 | 0.88 | 0.88 | 0.84 | 0.88 | 0.71 | 0.65 | 0.71 |
| A-890k-Wav | 0.00 | 0.00 | 0.01 | 0.08 | 0.05 | 0.31 | 0.31 | 0.30 | 0.31 | 0.36 | 0.34 | 0.36 |
| Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 0 | 2 | 0 |
| Média Postos | 2,19 | 2,28 | 1,53 | 2,03 | 2,78 | 1,19 | 1,69 | 2,61 | 1,69 | 1,61 | 2,78 | 1,61 |
| Posição | 2 | 3 | 1 | 2 | 3 | 1 | 1,5 | 3 | 1,5 | 1,5 | 3 | 1,5 |

Fonte: O AUTOR (2020).

C.3 RESUMO DOS TESTES DE HIPÓTESES

Tabela 38 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Detector | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1 | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2 | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3 | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1 | | | | | | | * | H ₂ | H ₀ | | | |
| ALOC2 | | | | | | | H ₁ | * | H ₁ | | | |
| ALOC3 | | | | | | | H ₀ | H ₂ | * | | | |
| eDetC1 | | | | | | | | | | * | H ₁ | H ₀ |
| eDetC2 | | | | | | | | | | H ₂ | * | H ₂ |
| eDetC3 | | | | | | | | | | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| $\sum H_1$ - Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 2 | 0 |
| $\sum H_2$ - Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 1 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | 2 | 0 | -2 | 2 | 1 | -2 | 1 | -1 | 2 | -1 |

Fonte: O AUTOR (2020).

Tabela 39 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Detector | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1 | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2 | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3 | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1 | | | | | | | * | H ₀ | H ₀ | | | |
| ALOC2 | | | | | | | H ₀ | * | H ₀ | | | |
| ALOC3 | | | | | | | H ₀ | H ₀ | * | | | |
| eDetC1 | | | | | | | | | | * | H ₁ | H ₀ |
| eDetC2 | | | | | | | | | | H ₂ | * | H ₂ |
| eDetC3 | | | | | | | | | | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 0 | 1 |
| $\sum H_1$ - Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| $\sum H_2$ - Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | 2 | 0 | -2 | 2 | 0 | 0 | 0 | -1 | 2 | -1 |

Fonte: O AUTOR (2020).

Tabela 40 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|--------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1a | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2a | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3a | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1a | | | | | | | * | H ₀ | H ₀ | | | |
| ALOC2a | | | | | | | H ₀ | * | H ₀ | | | |
| ALOC3a | | | | | | | H ₀ | H ₀ | * | | | |
| eDetC1a | | | | | | | | | | * | H ₁ | H ₀ |
| eDetC2a | | | | | | | | | | H ₂ | * | H ₂ |
| eDetC3a | | | | | | | | | | H ₀ | H ₁ | * |
| H ₀ | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 0 | 1 |
| H ₁ | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| H ₂ | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| H ₁ -H ₂ | 0 | -2 | 2 | 0 | -2 | 2 | 0 | 0 | 0 | -1 | 2 | -1 |

Fonte: O AUTOR (2020).

Tabela 41 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1 | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2 | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3 | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1 | | | | | | | * | H ₀ | H ₀ | | | |
| ALOC2 | | | | | | | H ₀ | * | H ₀ | | | |
| ALOC3 | | | | | | | H ₀ | H ₀ | * | | | |
| eDetC1 | | | | | | | | | | * | H ₁ | H ₀ |
| eDetC2 | | | | | | | | | | H ₂ | * | H ₂ |
| eDetC3 | | | | | | | | | | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | 0 | 1 |
| $\sum H_1$ - Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 |
| $\sum H_2$ - Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | 2 | 0 | -2 | 2 | 0 | 0 | 0 | -1 | 2 | -1 |

Fonte: O AUTOR (2020).

Tabela 42 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1 | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2 | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3 | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1 | | | | | | | * | H ₂ | H ₀ | | | |
| ALOC2 | | | | | | | H ₁ | * | H ₁ | | | |
| ALOC3 | | | | | | | H ₀ | H ₂ | * | | | |
| eDetC1 | | | | | | | | | | * | H ₂ | H ₀ |
| eDetC2 | | | | | | | | | | H ₁ | * | H ₁ |
| eDetC3 | | | | | | | | | | H ₀ | H ₂ | * |
| $\sum H_0$ - Empates | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| $\sum H_1$ - Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| $\sum H_2$ - Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 0 | 2 | 0 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | 2 | 0 | -2 | 2 | 1 | -2 | 1 | 1 | -2 | 1 |

Fonte: O AUTOR (2020).

Tabela 43 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos de múltipla análise dos trabalhos relacionados.

| Dataset | ADC1 | ADC2 | ADC3 | ALHDC1 | ALHDC2 | ALHDC3 | ALOC1 | ALOC2 | ALOC3 | eDetC1 | eDetC2 | eDetC3 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ADC1 | * | H ₂ | H ₁ | | | | | | | | | |
| ADC2 | H ₁ | * | H ₁ | | | | | | | | | |
| ADC3 | H ₂ | H ₂ | * | | | | | | | | | |
| ALHDC1 | | | | * | H ₂ | H ₁ | | | | | | |
| ALHDC2 | | | | H ₁ | * | H ₁ | | | | | | |
| ALHDC3 | | | | H ₂ | H ₂ | * | | | | | | |
| ALOC1 | | | | | | | * | H ₂ | H ₀ | | | |
| ALOC2 | | | | | | | H ₁ | * | H ₁ | | | |
| ALOC3 | | | | | | | H ₀ | H ₂ | * | | | |
| eDetC1 | | | | | | | | | | * | H ₂ | H ₀ |
| eDetC2 | | | | | | | | | | H ₁ | * | H ₁ |
| eDetC3 | | | | | | | | | | H ₀ | H ₂ | * |
| $\sum H_0$ - Empates | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| $\sum H_1$ - Vitórias | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| $\sum H_2$ - Derrotas | 1 | 2 | 0 | 1 | 2 | 0 | 0 | 2 | 0 | 0 | 2 | 0 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | 2 | 0 | -2 | 2 | 1 | -2 | 1 | 1 | -2 | 1 |

Fonte: O AUTOR (2020).

C.4 CONSIDERAÇÕES FINAIS

O objetivo deste capítulo foi classificar o desempenho dos métodos de comitês avaliados em termos estatísticos usando teste de hipótese e média dos postos. Resultados experimentais mostram os melhores comitês por método usando a média geral na classificação, considerando os três critérios de avaliação. Por fim, os resultados de desempenhos dos métodos foi calculado com base na metodologia de avaliação proposta nesta tese (apresentada na seção 3) que classifica o desempenho com base nas diferenças estatística calculadas entre pares de tratamento do grupo de tratamentos.

A análise dos resultados indicou que há uma variação significativa entre desempenho dos diferentes métodos e três comitês selecionados. Considerando como parâmetros de bom e ruim os métodos de melhor e pior desempenho, a diferença entre os resultados passa dos 100% entre os valores absolutos alcançados em relação as médias obtidas pelos testes de hipóteses. Como exemplo, o método AD com o comitê ADC₃ obteve rank 2 enquanto que com o comitê AD.C₂ foi -2.

Vale ressaltar que não foi o objetivo principal da avaliação realizada nesta seção identificar o método que obteve melhor desempenho. Foram utilizadas três parametrizações para cada um dos métodos AD, ALHD, ALO e *e*-Detector, objetivando avaliar qual delas proporcionariam ao método melhores resultados. A avaliação completa entre o desempenho dos métodos será feita na próxima seção (seção 4), incluindo também os métodos propostos. Assim sendo, na próxima seção os métodos propostos serão apresentados e avaliados, seguindo a seguinte ordem: 1) análise individual dos métodos propostos com o objetivo de encontrar as melhores parametrizações; e 2) as melhores parametrizações serão utilizadas para comparar os métodos e parametrizações desta seção, juntamente com os detectores membros utilizados para formar os comitês.

APÊNDICE D – RESUMO DOS TESTES COM COMITÊS

D.1 DDE, MSDDE, MDDE RESULTADOS E DISCUSSÃO

D.1.1 DDE - Resultados e Discussão

As Tabelas 44 e 45 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê que obteve os melhores resultados com o HT foi DDE.C_{2b}, com 3,13. Já com o NB foram os comitês DDE.C_{1a} e DDE.C_{3a}, ambos com 3,27. Em relação aos tipos de graduação, tanto para os resultados de HT quanto de NB os métodos obtiveram melhores resultados para bases abruptas. Dentre as bases reais o algoritmo que obteve melhor resultado foi o DDE.C_{2a} para o HT e NB.

A Figura 41 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 62 e 63. Em termos das hipóteses testadas, o comitê DDE.C_{2b} com HT forneceu o maior número de vitórias (7) e em termos estatísticos ele foi superior estatisticamente a todas as variações. Já com NB os melhores em termos de quantidade de vitórias foram o DDE.C_{1a}, DDE.C_{2a}, DDE.C_{2b} e DDE.C_{3a}, ambos com 4 vitórias e em termos estatísticos foram iguais. O comitê com maior número de derrotas usando o HT e NB foi DDE.C_{2c}, com 3 e 7 derrotas, respectivamente.

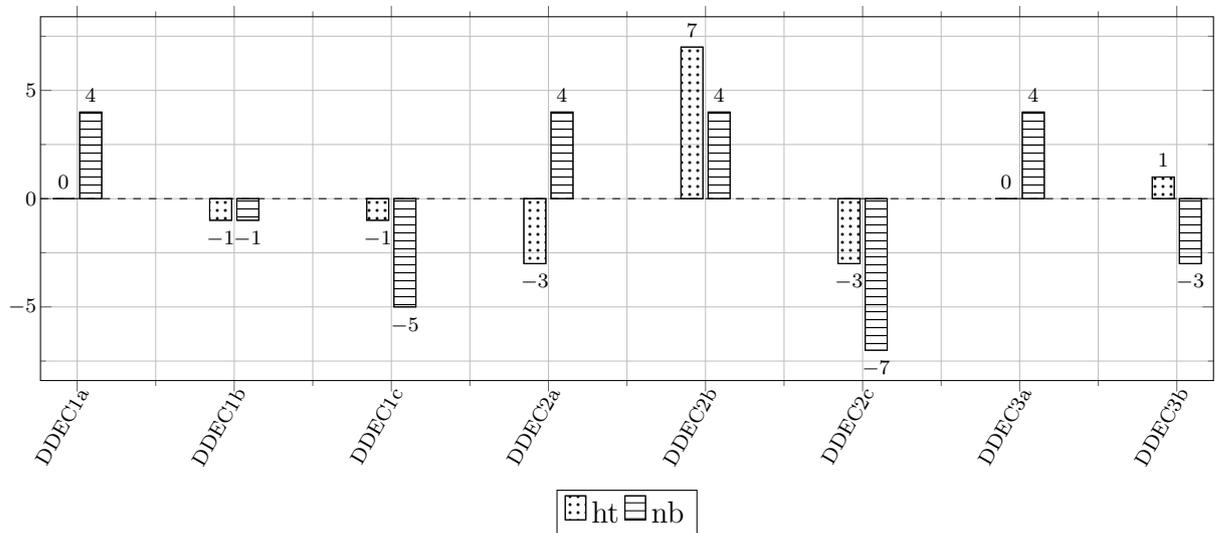
O recurso do mapa de calor contribui para contabilizar os métodos mais eficientes. Como pode ser visto com o HT, o comitê DDE.C_{2b} foi o método que menos vezes obteve resultados ruins, ou seja, ele apresentou apenas em três ocasiões a célula não verde. Note-se que com o NB não foi muito diferente com os comitês DDE.C_{1a} e DDE.C_{3a}, ambos apresentaram o menor número de vezes a célula não verde (6).

Os melhores resultados foram obtidos pelos comitês DDE.C_{1a}, DDE.C_{2a}, DDE.C_{2b} e DDE.C_{3a}, todos com diferenças calculadas acima de zero, como pode ser visto na Figura 41 (b). Apesar disso, é importante destacar o desempenho do DDE.C_{2b} foi sólido na maioria das situações e houve superioridade estatística em relação aos demais comitês com o HT. A Figura 41 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB, logo é possível observar que o comitê DDE.C_{2b} obteve a maior média (5,5) no desempenho acumulado, seguidos pelos comitês DDE.C_{1a}, DDE.C_{3a} empatados com 2 e DDE.C_{2a} com 0,5. Vale ressaltar que metade do número de comitês obtiveram rank com valor negativo (abaixo de zero).

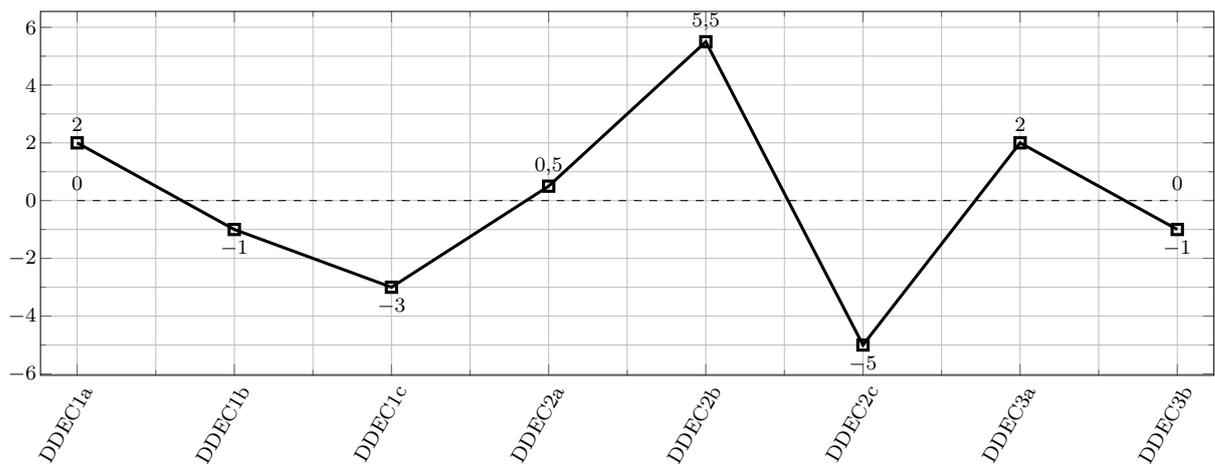
O comitê DDE.C_{1c} com HT obteve bons resultados nas bases com 890k e ruins nas base com 20k. No entanto, o comitê DDE.C_{2c} não acompanhou tal desempenho, apesar de também possuir três detectores. Já nas bases com NB, ambos apresentaram os piores resultados.

Figura 41 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do DDE.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



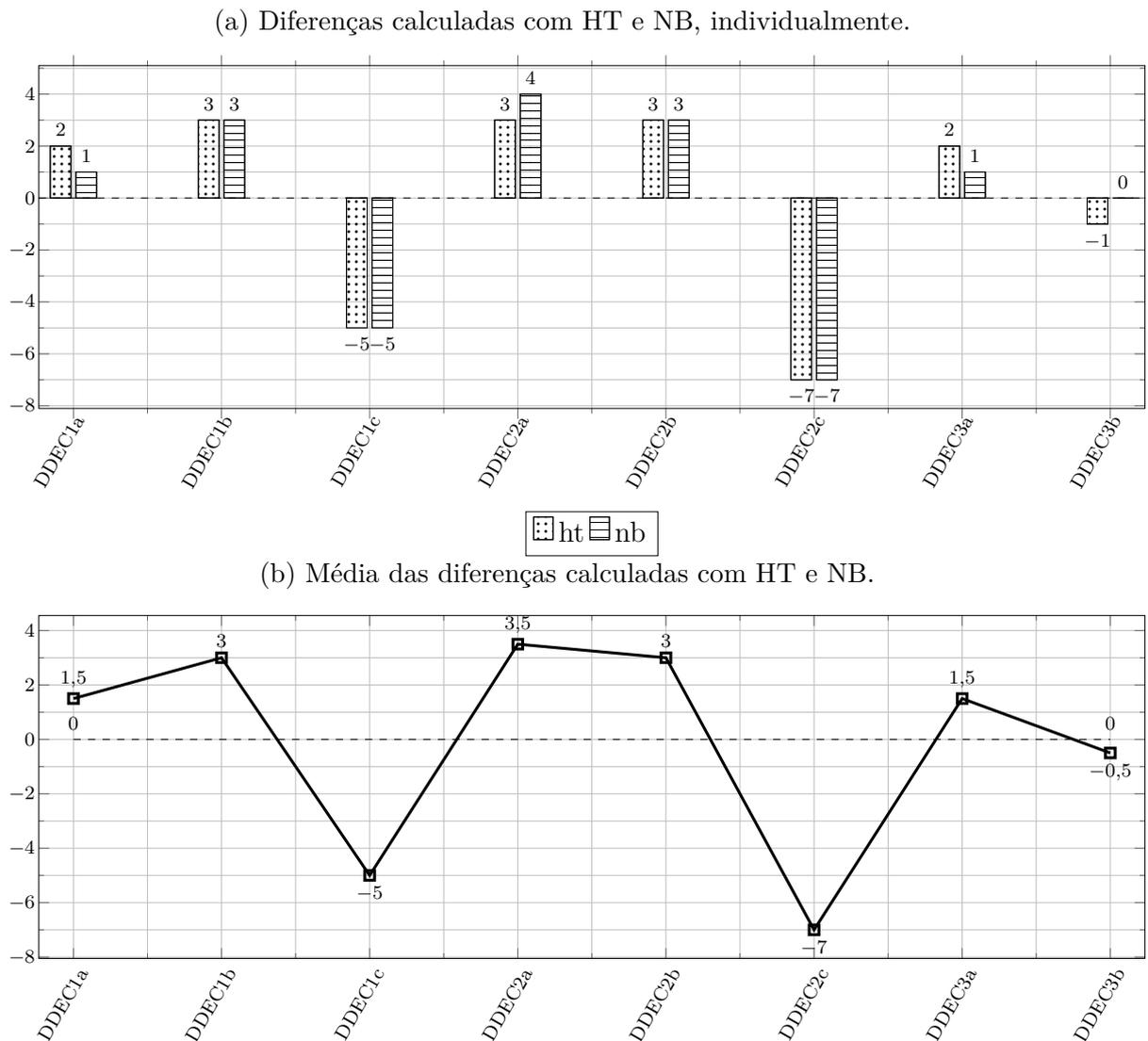
Fonte: O AUTOR (2020).

As Tabelas 46 e 47 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê DDE.C_{2a} obteve os melhores resultados com o HT (3,25) e NB (2,97). Assim como o DDE.C_{1c} obteve os piores resultados com o HT e NB. Como pode se ver com o HT e NB, o comitê DDE.C_{2a} foi o método que mais vezes obteve os melhores resultados nas bases abruptas. Note-se que a eficiência dos métodos cai bastante nas bases graduais, apesar disso o comitê DDE.C_{2a} conseguiu manteve-se com o melhor resultado geral.

A Figura 42 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 64 e 65. Em termos das hipóteses testadas, os comitês

DDE.C_{1b}, DDE.C_{2a} e DDE.C_{2c} com HT forneceram os melhores desempenhos, ambos com 3 vitórias e estatisticamente iguais. Tratando-se do NB, o comitê DDE.C_{2a} obteve 4 vitórias enquanto o DDE.C_{1b} e o DDE.C_{2c} obtiveram 3 vitórias. Apesar disso, eles foram iguais estatisticamente.

Figura 42 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do DDE.



Fonte: O AUTOR (2020).

Os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês DDE.C_{1a}, DDE.C_{1b}, DDE.C_{2a}, DDE.C_{2b}, DDE.C_{3a} e DDE.C_{3b}, todos com rank de avaliação acima de zero, mostrados na Figura 42 (b). Apesar disso, vale a pena destacar que o desempenho de DDE.C_{1b}, DDE.C_{2a} e DDE.C_{2b} foram bons na maioria das situações com o HT e com o NB. A Figura 42 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que o DDE.C_{2a} obteve a maior média (3,5) no desempenho acumulado, seguido pelos comitês DDE.C_{1b}, DDE.C_{2b}

(ambos com 3) DDE.C_{1a} e DDE.C_{3a} (com 1, 5). Os demais obtiveram média negativa no saldo das diferenças entre vitórias e derrotas.

As Tabelas 48 e 49 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê DDE.C_{2b} obteve os melhores resultados com o HT (2, 72). Já com o NB, o melhor comitê foi o DDE.C_{3b} com 2, 39.

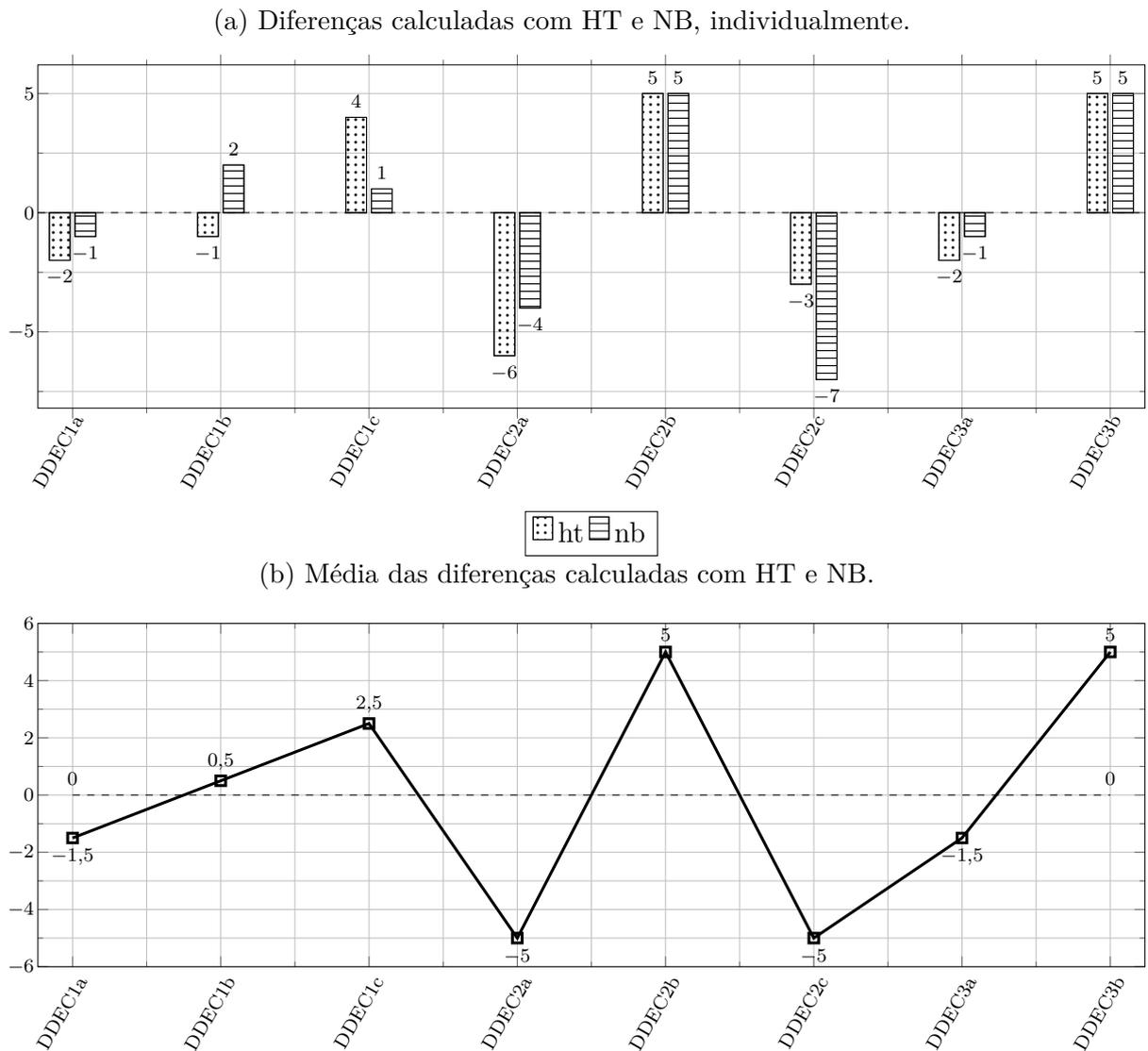
A Figura 43 (a) mostra o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 66 e 67. Em termos das hipóteses testadas, os comitês DDE.C_{2b} e DDE.C_{3b} forneceram os maiores números de superioridades estatísticas com HT e com NB, ambos com 5 vitórias. Entretanto, o comitê DDE.C_{1c} foi igual estatisticamente aos comitês DDE.C_{2b} e DDE.C_{3b} com o HT e NB. Os comitês com maiores números de derrotas usando HT e NB foram DDE.C_{2a} e DDE.C_{2c}, respectivamente. Logo, tratando-se da ineficiência, eles foram os piores comitês avaliados em termos absolutos e em termos estatísticos.

Os melhores resultados foram obtidos pelos comitês DDE.C_{1b}, DDE.C_{1c}, DDE.C_{2b}, e DDE.C_{3b}, todos com avaliação acima de zero, mostrados na Figura 43 (b). Apesar disso, vale a pena destacar que o desempenho dos comitês DDE.C_{2b} e DDE.C_{3b} foram sólidos na maioria das situações e líderes em vitórias com HT e NB. A Figura 43 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que os comitês DDE.C_{2b} e DDE.C_{3b} obtiveram as maiores médias no desempenho acumulado, seguidos pelo comitê DDE.C_{1c} (2, 5) e DDE.C_{1b} (com 0, 5). Os demais comitês ficaram com média negativa em número de vitórias, ou seja, foram inferiores mais vezes que superiores estatisticamente.

A Figura 44 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 44 (a) é feito o rank da média das diferenças. A Figura 44 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 44 (a) ou no da Figura 44 (b), o detector de maior valor absoluto foi o DDE.C_{2b} com médias 4, 5 e 7, 17, respectivamente. Vale ressaltar que por meio do rank da média (Figura 44 (a)) é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários. O comitê DDE.C_{2c} apresentou os piores resultados.

O DDE.C_{2b} foi o único comitê no qual em todas as avaliações não obteve derrota. Além disso, ele também foi o comitê de melhor classificação ordenada por cálculo de média dos ranks e por cálculo de média dos postos. Assim, o DDE.C_{2b} apresenta-se como o melhor comitê entre os comitês avaliados. Vale ressaltar que o DDE.C_{2b} (4, 5) obteve uma média das diferenças mais que três vezes maior que ao segundo colocado DDE.C_{3b} com 1, 17.

Figura 43 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do DDE.



Fonte: O AUTOR (2020).

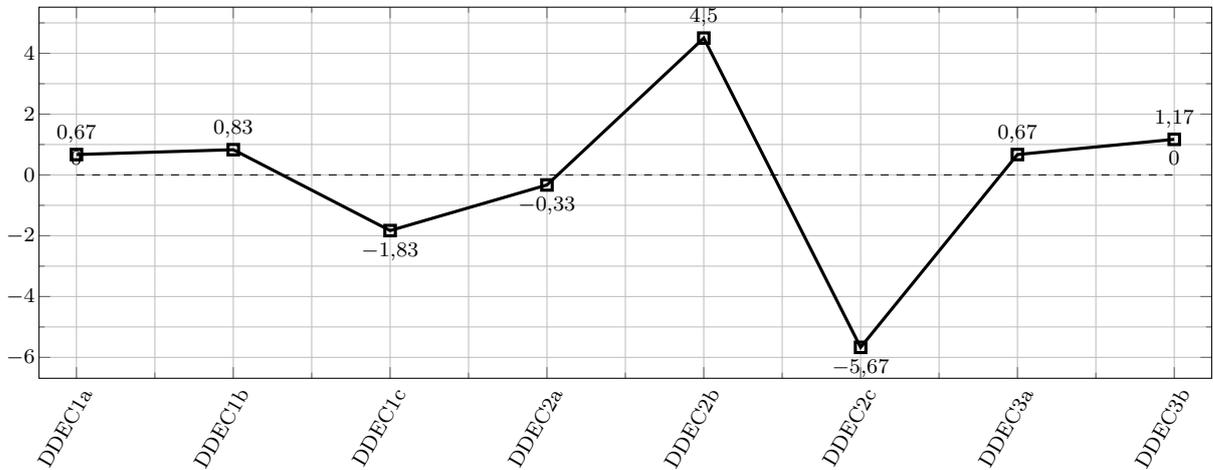
D.1.2 MSDDE - Resultados e Discussão

As Tabelas 50 e 51 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê que obteve os melhores resultados com o HT e com o NB foi MSDDE.C₂b.

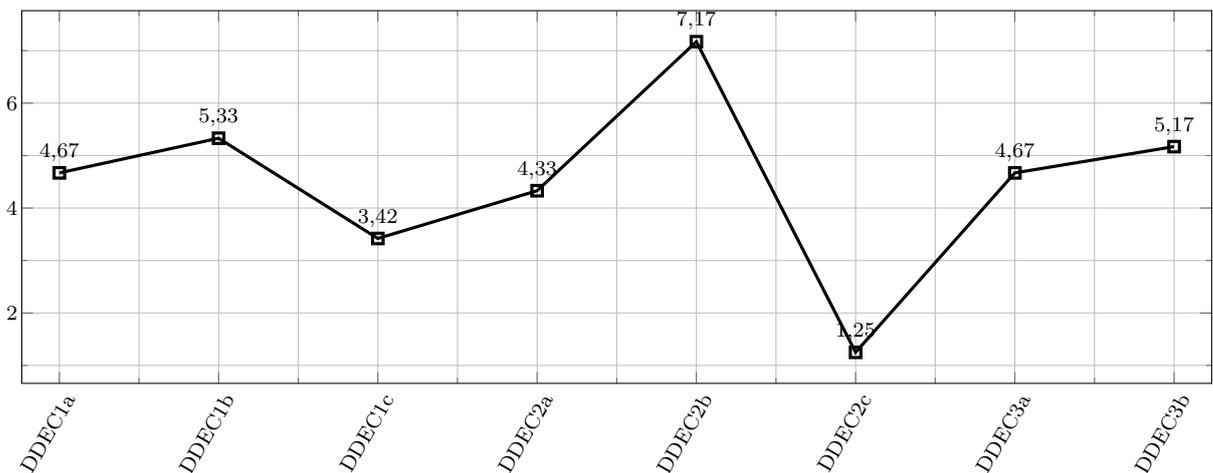
A Figura 45 (a) traz o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 68 e 69. Em termos das hipóteses testadas, o comitê MSDDE.C₂b com HT forneceu o maior número de vitórias (7) e em termos estatísticos ele foi superior estatisticamente a todas as variações. Já com NB, o melhor em termos de quantidade de vitórias foi o MSDDE.C₁b (5). Apesar disso, ele obteve resultados

Figura 44 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do DDE.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.

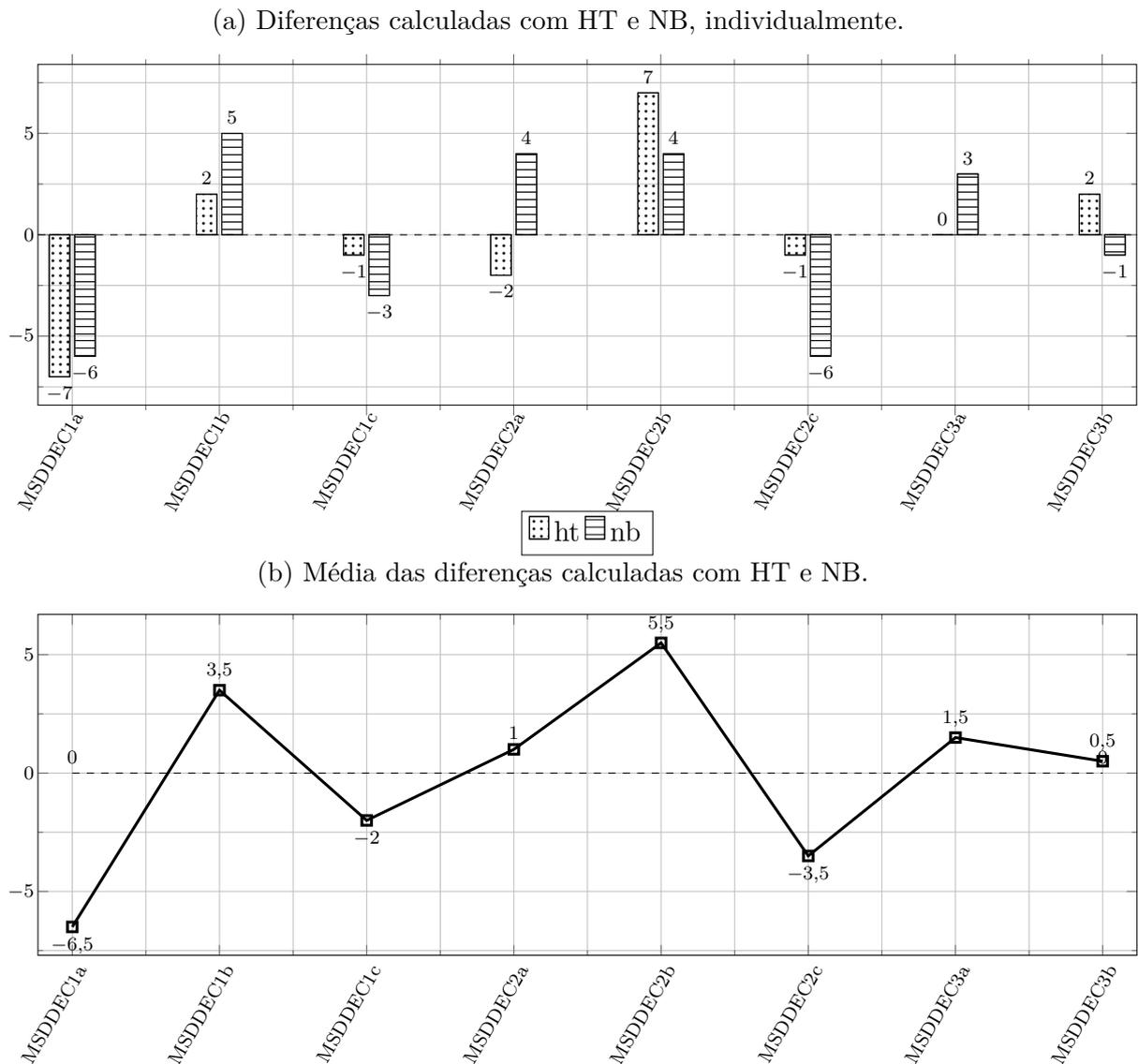


Fonte: O AUTOR (2020).

iguais estatisticamente aos comitês MSDDE.C_{2a} e MSDDE.C_{2b}. Os comitês com maiores números de derrotas foram: MSDDE.C_{1a}, utilizando o classificador HT (7 derrotas); e MSDDE.C_{1a} e MSDDE.C_{2c}, utilizando o classificador NB (6 derrotas).

Os melhores resultados foram obtidos pelos comitês MSDDE.C_{1b}, MSDDE.C_{2a}, MSDDE.C_{2b}, MSDDE.C_{3a} e MSDDE.C_{3b}, todos com diferenças calculadas acima de zero, como pode ser visto na Figura 45 (b). Apesar disso, vale a pena destacar que o desempenho do MSDDE.C_{2b} foi sólido na maioria das situações e houve superioridade estatística em relação aos demais comitês com o HT. A Figura 45 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que o comitê MSDDE.C_{2b} obteve a maior média (5,5) no desempenho acumulado, seguidos pelo comitê MSDDE.C_{1b} com 3,5. Vale ressaltar que menos da metade do número de comitês obtiveram rank com valor negativo (abaixo de zero).

Figura 45 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do MSDDE.



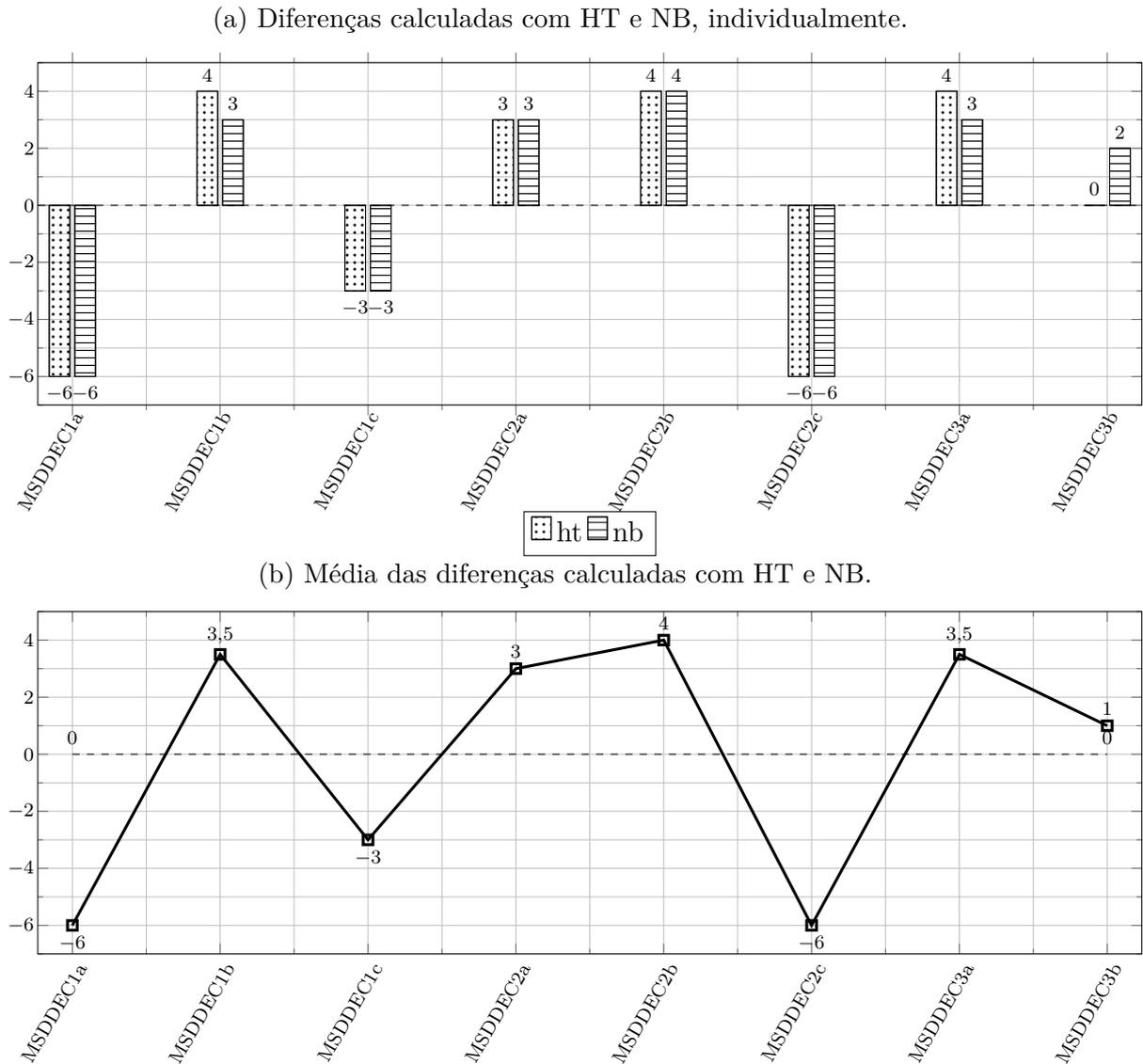
Fonte: O AUTOR (2020).

As Tabelas 52 e 53 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê MSDDE.C₂b obteve os melhores resultados com o HT (3, 11) e NB (3, 17).

A Figura 46 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 70 e 71. Em termos das hipóteses testadas, os comitês MSDDE.C₁b e MSDDE.C₃a com HT forneceram os melhores desempenhos, ambos com 4 vitórias e foram estatisticamente iguais. Tratando-se do NB, o comitê MSDDE.C₂b obteve 4 vitórias, enquanto o MSDDE.C₁b, MSDDE.C₂a e MSDDE.C₃a obtiveram 3

vitórias. Apesar disso, eles (os quatro) foram iguais estatisticamente.

Figura 46 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do MSDDE.



Fonte: O AUTOR (2020).

De acordo com a Figura 46 (b), os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês MSDDE.C₁b, MSDDE.C₂a, MSDDE.C₂b, MSDDE.C₃a e MSDDE.C₃b, todos com rank de avaliação acima de zero. Apesar disso, vale a pena destacar que o desempenho de MSDDE.C₁b, MSDDE.C₂a, MSDDE.C₂b e MSDDE.C₃a foram bons na maioria das situações com o HT e com o NB. A Figura 46 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que o MSDDE.C₂b obteve a maior média (4) no desempenho acumulado, seguido pelos comitês MSDDE.C₁b e MSDDE.C₃a, ambos com 3,5, MSDDE.C₂a com 3 e MSDDE.C₃b com 1, ficando os

demais com média negativa no saldo das diferenças entre vitórias e derrotas.

As Tabelas 54 e 55 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês MSDDE.C₂b com 2,50 e MSDDE.C₃b com 2,33 obtiveram os melhores resultados com HT e NB, respectivamente.

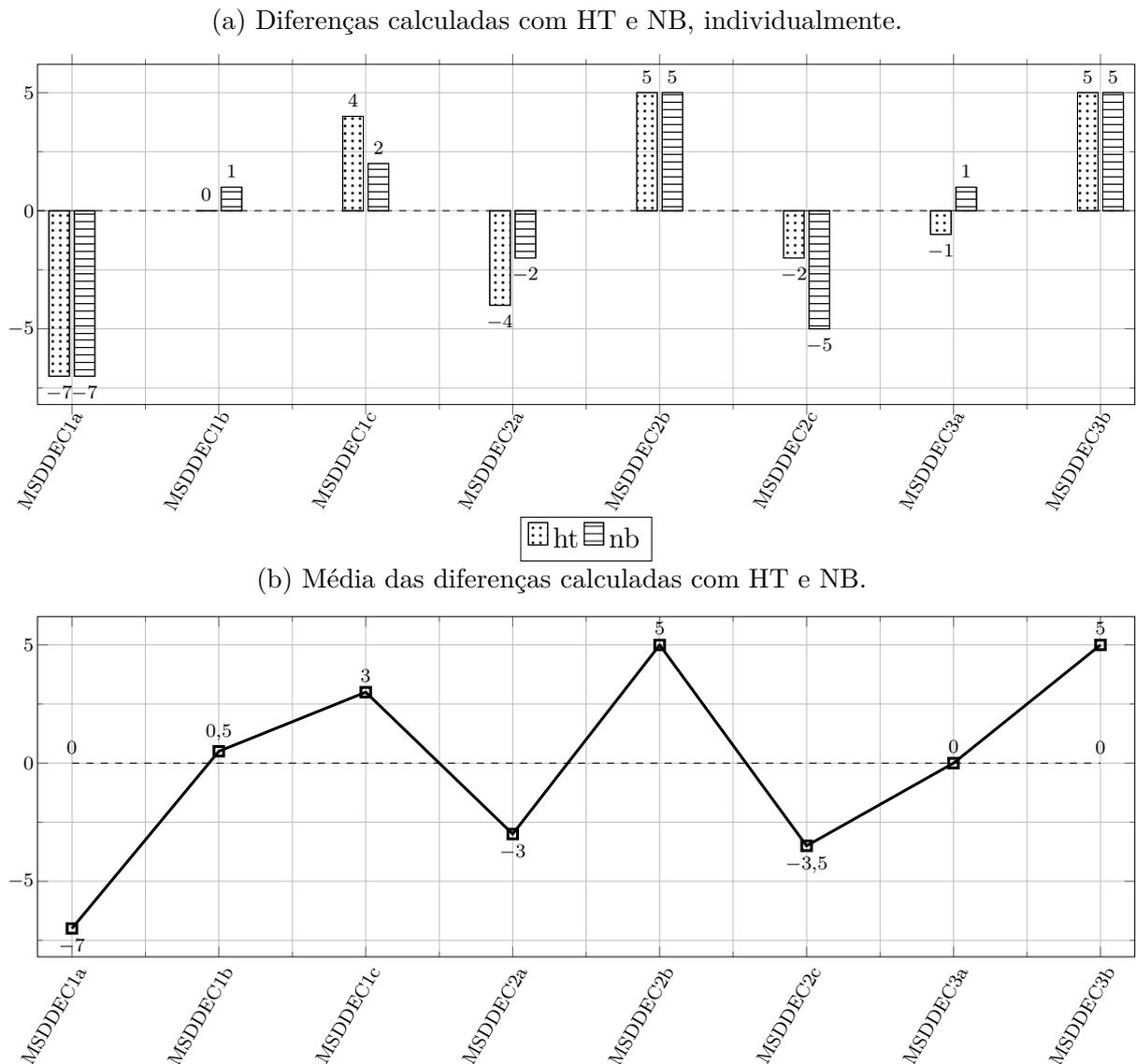
A Figura 47 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 72 e 73. Em termos das hipóteses testadas, os comitês MSDDE.C₂b e MSDDE.C₃b forneceram os maiores números de superioridades estatísticas com HT e com NB, ambos com 5 vitórias. Entretanto, o comitê MSDDE.C₁c foi igual estatisticamente aos MSDDE.C₂b e MSDDE.C₃b com o HT e NB. O comitê com maiores números de derrotas usando HT e NB foi o MSDDE.C₁a com 7 derrotas. Logo, tratando-se de ineficiência, eles foram em termos absolutos e em termos estatísticos os piores comitês avaliados.

Os melhores resultados foram obtidos pelos comitês MSDDE.C₁b, MSDDE.C₁c, MSDDE.C₂b, MSDDE.C₃a e MSDDE.C₃b, todos com avaliação igual ou acima de zero, mostrados na Figura 47 (b). Apesar disso, vale a pena destacar que o desempenho dos comitês MSDDE.C₂b e MSDDE.C₃b foram sólidos na maioria das situações e líderes em vitórias com HT e NB. A Figura 47 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que os comitês MSDDE.C₂b e MSDDE.C₃b obtiveram as maiores médias no desempenho acumulado, seguidos pelos comitês MSDDE.C₁c (3), MSDDE.C₁b com 0,5 e MSDDE.C₁c com 0. Os demais comitês ficaram com média negativa em número de vitórias, ou seja, foram inferiores mais vezes que superiores estatisticamente.

A Figura 48 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 48 (a) é feito o rank da média das diferenças. A Figura 48 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 48 (a) ou no da Figura 48 (b), o detector de maior valor absoluto foi o MSDDE.C₂b com médias 4,83 e 7,82, respectivamente. Vale ressaltar que por meio do rank da média (Figura 48 (a)) é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários. O comitê MSDDE.C₁a apresentou os piores resultados.

O MSDDE.C₂b foi o único comitê no qual em todas as avaliações não obteve derrota. Além disso, ele também foi o comitê de melhor classificação ordenada por cálculo de média dos ranks e por cálculo de média dos postos. Assim, o MSDDE.C₂b apresenta-se como o melhor comitê entre os comitês avaliados. Vale ressaltar que o MSDDE.C₂b (4,83) obteve uma média das diferenças quase duas vezes maior que ao segundo colocado MSDDE.C₁b,

Figura 47 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do MSDDE.



Fonte: O AUTOR (2020).

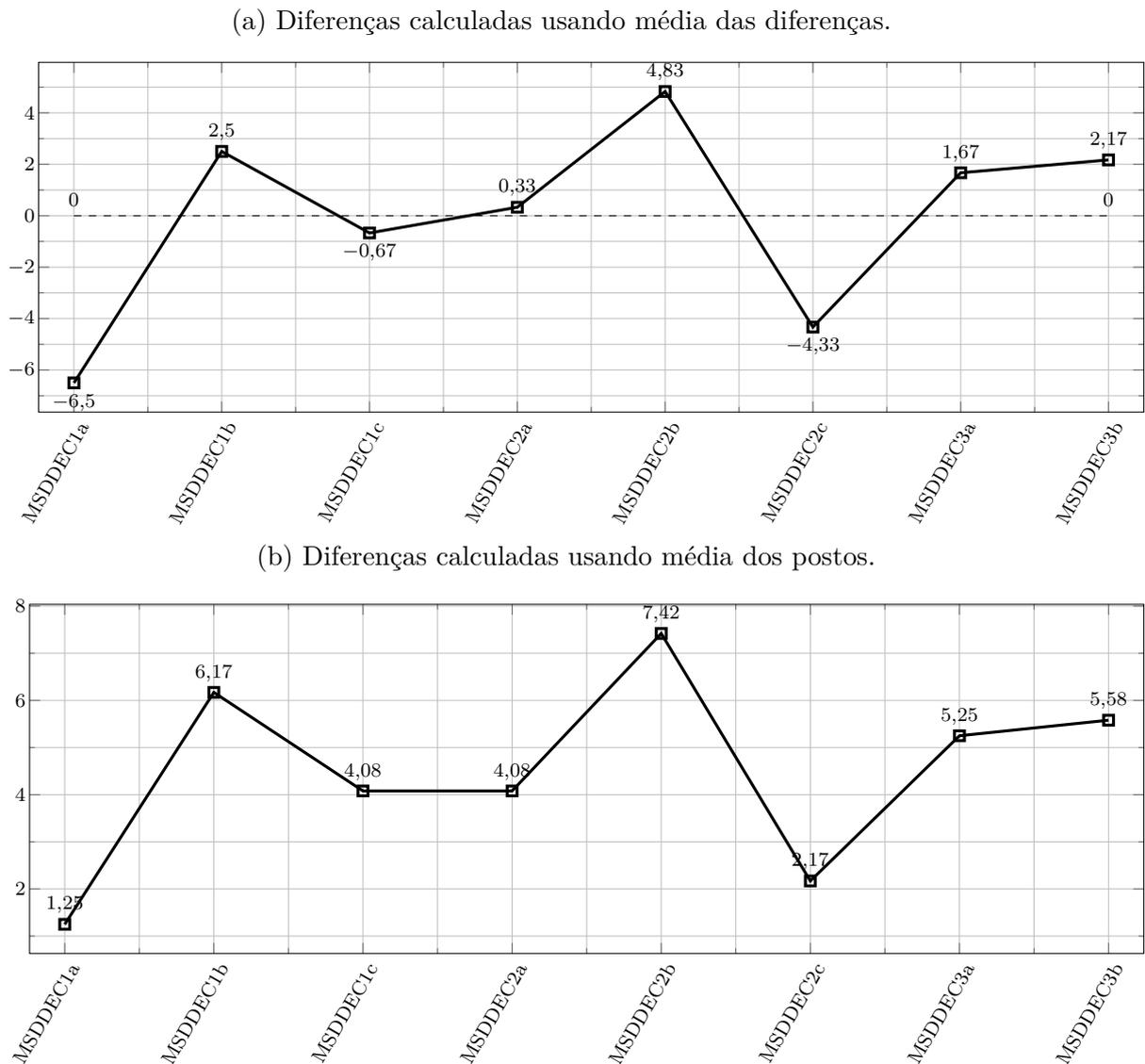
com 2, 5.

D.1.3 MDDE - Resultados e Discussão

As Tabelas 56 e 57 apresentam os resultados das médias de acurácias dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês que obtiveram os melhores resultados com o HT e NB foram MDDEC_{2a} (2, 44) e MDDEC_{2b} (1, 77), respectivamente.

A Figura 49 (a) apresenta o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 74 e 75. Em termos das hipóteses testadas, o comitê MDDEC_{2a}

Figura 48 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do MSDDE.

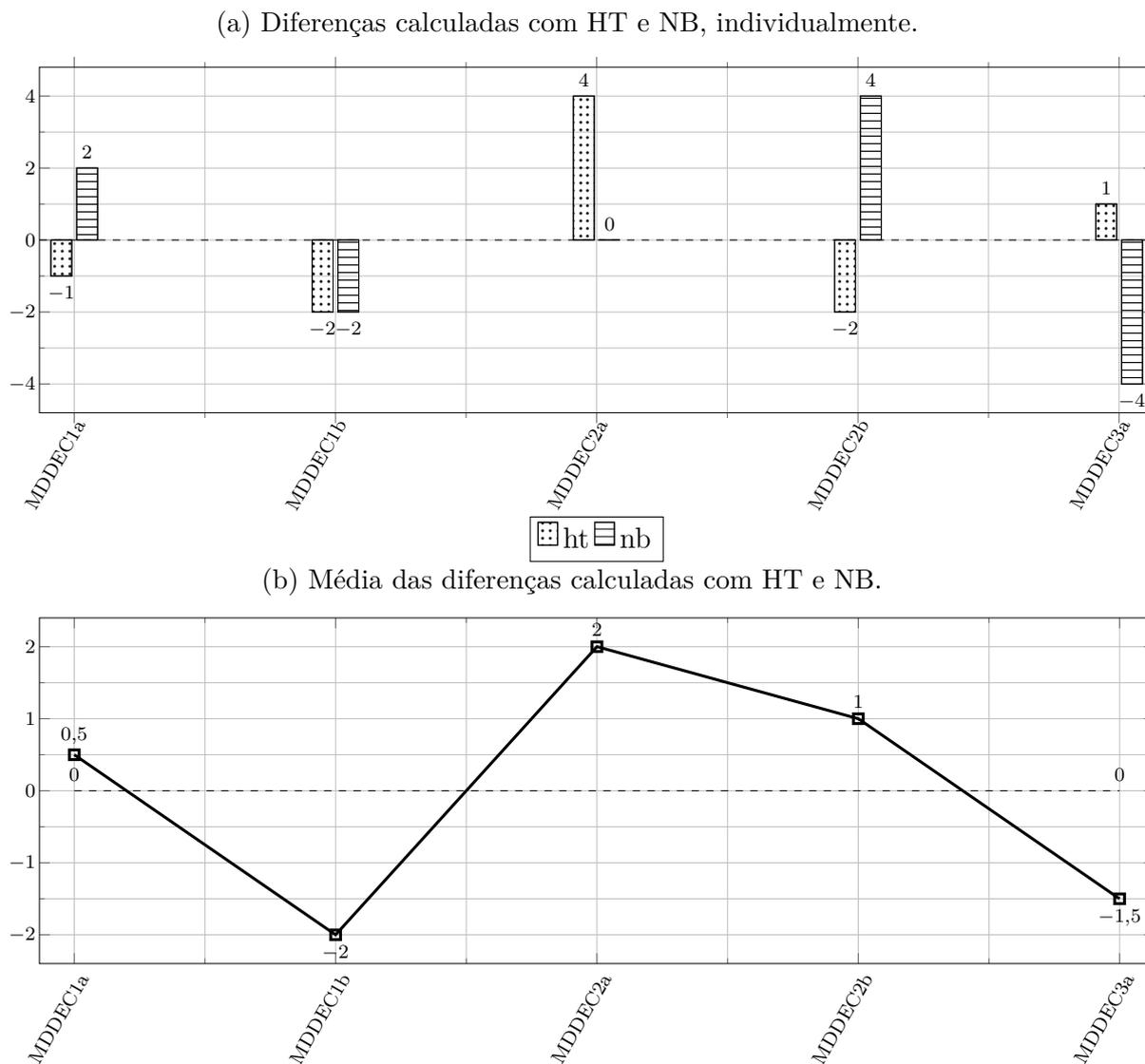


Fonte: O AUTOR (2020).

com HT forneceu o maior número de vitórias (4) e em termos estatísticos ele foi superior estatisticamente a todas as variações. Já com NB o melhor em termos de quantidade de vitórias foi o MDDEC_{2b} (4) e em termos estatísticos ele também foi superior estatisticamente a todas as variações. Os comitês com maiores números de derrotas foram MDDEC_{1b} e MDDEC_{2b} com o HT e 2 derrotas. Já com o NB foi o comitê MDDEC_{3a} com 4 derrotas.

É possível notar que os melhores resultados nos conjuntos de dados avaliados foram obtidos pelos comitês MDDEC_{1a}, MDDEC_{2a} e MDDEC_{2b}, todos com diferenças calculadas acima de zero, como pode ser visto na Figura 49 (b). Apesar disso, vale a pena destacar que o desempenho do método nos comitês MDDEC_{2a} e MDDEC_{2b} foram

Figura 49 – Resumo dos ranks das acurácias com HT e NB. Considerando comparações com parametrizações do MDDE.



Fonte: O AUTOR (2020).

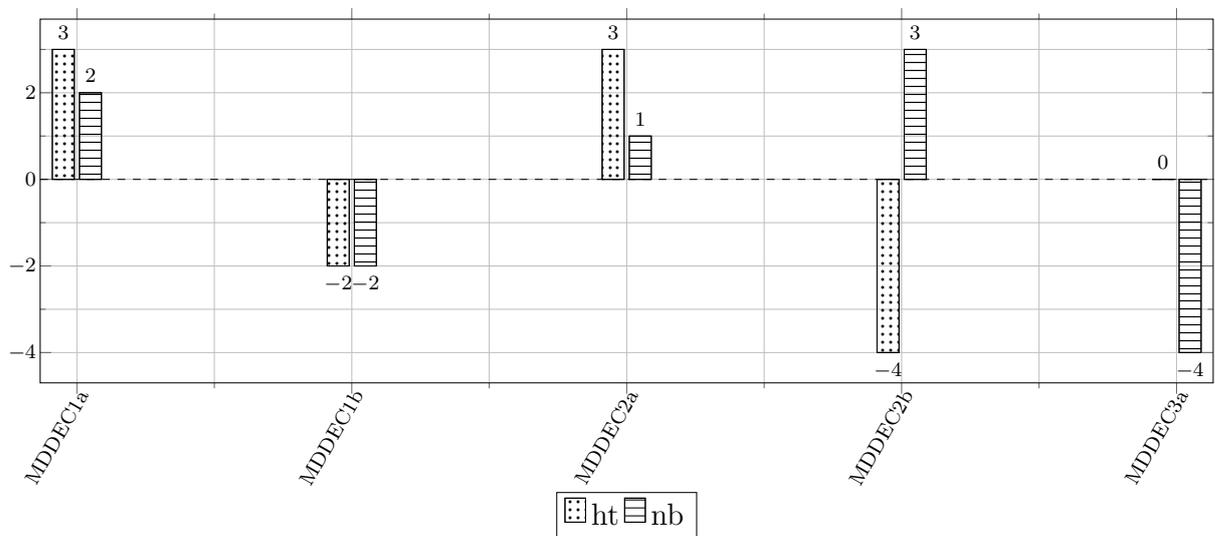
relevantes na maioria das situações e houve superioridade estatística em relação aos demais comitês com o HT e NB, respectivamente. A Figura 49 (b) mostra a diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo é possível observar que o comitê MDDEC_{2a} obteve a maior média (2) no desempenho acumulado, seguido pelo comitê MDDEC_{2b} com 1. Vale ressaltar que menos da metade do número de comitês obtiveram rank com valor negativo (abaixo de zero).

As Tabelas 58 e 59 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, os comitês MDDEC_{2a} e MDDEC_{2b} obtiveram os melhores resultados com o HT (2, 11) e NB (1, 97), respectivamente.

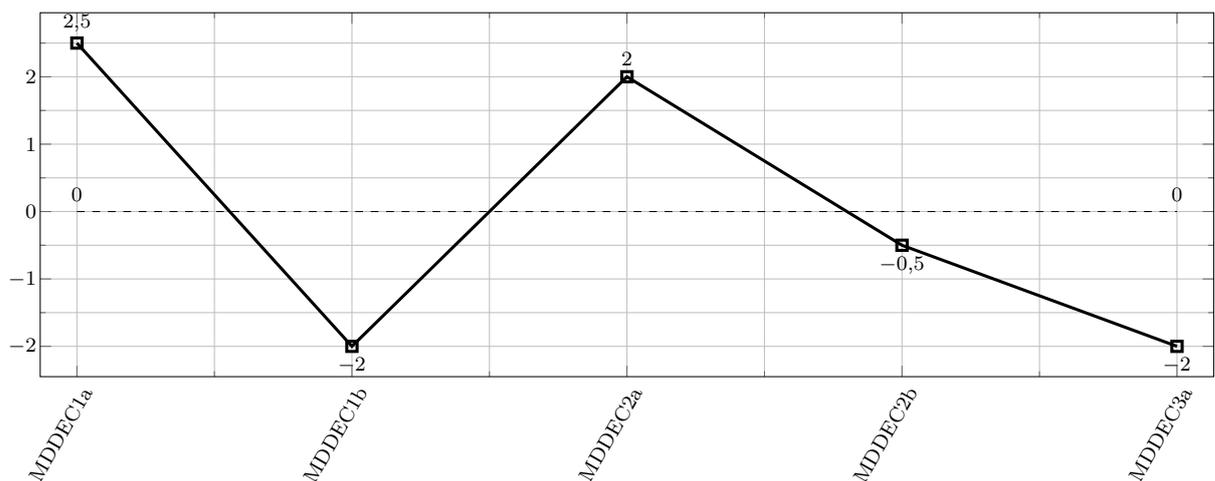
A Figura 50 (a) traz o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 76 e 77. Em termos das hipóteses testadas, os comitês MDDEC_{1a} e MDDEC_{2a} com o HT forneceram os melhores desempenhos, ambos com 3 vitórias e foram estatisticamente iguais e superiores aos demais. Tratando-se do NB, os comitês MDDEC_{1a} e MDDEC_{2a} também foram iguais estatisticamente. Apesar disso o número de vitórias do MDDEC_{2b} (3) foi maior que do MDDEC_{1a} (2).

Figura 50 – Resumo dos ranks dos MTRs com HT e NB. Considerando comparações com parametrizações do MDDE.

(a) Diferenças calculadas com HT e NB, individualmente.



(b) Média das diferenças calculadas com HT e NB.



Fonte: O AUTOR (2020).

Atende-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês MDDEC_{1a} e MDDEC_{2a}, ambos com rank de avaliação acima de zero, mostrados na Figura 50 (b). Apesar disso, vale a pena destacar que o desempenho deles foram bons na maioria das situações com o HT e com o NB. A Figura 50 (b) mostra a

diferença média calculada, levando em consideração as hipóteses calculadas com HT e NB. Logo, é possível observar que o MDDEC_{1a} obteve a maior média (2,5) no desempenho acumulado, seguido pelo comitê MDDEC_{2a}, com 2. Os demais ficaram com média negativa no saldo das diferenças entre vitórias e derrotas.

As Tabelas 60 e 61 apresentam os resultados das médias de dos métodos testados em todos os conjuntos de dados selecionados, bem como em suas classificações usando HT e NB, respectivamente. Em termos absolutos, tomando como base a média do cálculo dos postos, o comitê MDDEC_{2a} obteve os melhores resultados tanto com o HT (2,28) quanto com o NB (2,06).

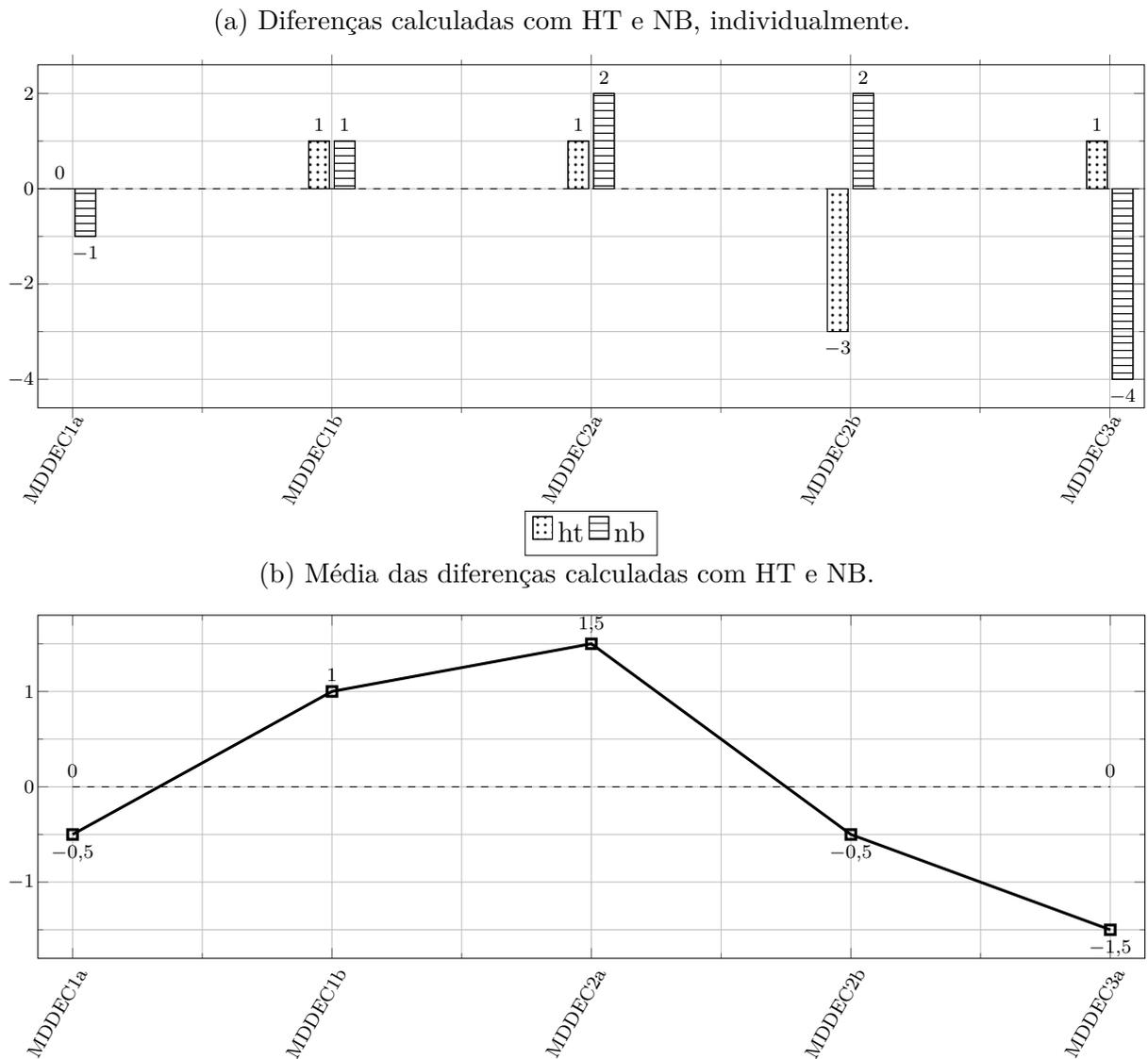
A Figura 51 (a) traz o resumo dos resultados com HT e NB apresentados no Apêndice D nas Tabelas 78 e 79. Em termos das hipóteses testadas, os comitês MDDEC_{1b}, MDDEC_{2a} e MDDEC_{3a} com o HT obtiveram 1 vitória. Já com o NB foram MDDEC_{2a} e MDDEC_{2b}, ambos com 2 vitórias. Entretanto, os comitês MDDEC_{3a} foi igual estatisticamente aos MDDEC_{1a}, MDDEC_{1b} e MDDEC_{2a} com o HT e com o NB os comitês MDDEC_{1b} e MDDEC_{2b} foram iguais ao MDDEC_{2a}. Os comitês com maiores números de derrotas usando HT e NB foram MDDEC_{2b} (3) e MDDEC_{3a} (4), respectivamente. Logo, tratando-se da ineficiência, eles foram em termos absolutos e em termos estatísticos os piores comitês avaliados.

Nota-se que os melhores resultados nos conjuntos de dados avaliados são obtidos pelos comitês MDDEC_{1b} e MDDEC_{2a}, todos com avaliação igual ou acima de zero, mostrados na Figura 51 (b). Apesar disso, vale a pena destacar que o desempenho dos comitês MDDEC_{2b} e MDDEC_{3a} foram sólidos na maioria das situações e líderes em vitórias com HT e NB, respectivamente. A Figura 51 (b) mostra a diferença média, levando em consideração as hipóteses calculadas com HT e NB. Assim, é possível observar que o comitê MDDEC_{2a} obteve a maior média no desempenho acumulado, seguidos pelo comitê MDDEC_{1b} (1). Os demais comitês ficaram com média negativa em número de vitórias, ou seja, foram inferiores estatisticamente mais vezes do que superiores.

A Figura 52 apresenta os ranks obtidos pelos detectores levando em consideração os três critérios avaliados (acurácia, MTR e MCC) e os dois classificadores (HT e NB). Na Figura 52 (a) é feito o rank da média das diferenças. A Figura 52 (b), por outro lado, apresenta o rank da média do cálculo dos postos. Em ambos os ranks, seja no da Figura 52 (a) ou no da Figura 52 (b), o detector de maior valor absoluto foi o MDDEC_{2a} com médias 1,83 e 4, respectivamente. Vale ressaltar que por meio do rank da média (Figura 52 (a)) é possível visualizar os números negativos que mostram quais detectores obtiveram mais derrotas que vitórias em termos estatísticos, sendo considerado um detector menos eficiente nesses cenários. O comitê MDDEC_{3a} apresentou os piores resultados.

O MDDEC_{2a} foi bem na maioria das avaliações, sendo inferior estatisticamente a outros comitês apenas nos critérios acurácia e MTR com o NB. Em termos gerais, ele foi

Figura 51 – Resumo dos ranks dos MCCs com HT e NB. Considerando comparações com parametrizações do MDDE.

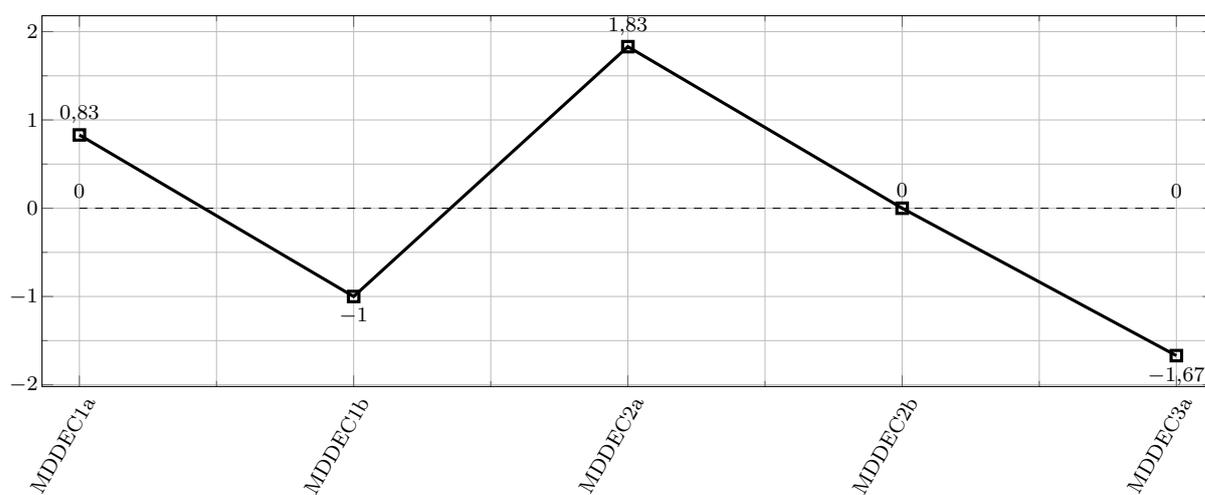


Fonte: O AUTOR (2020).

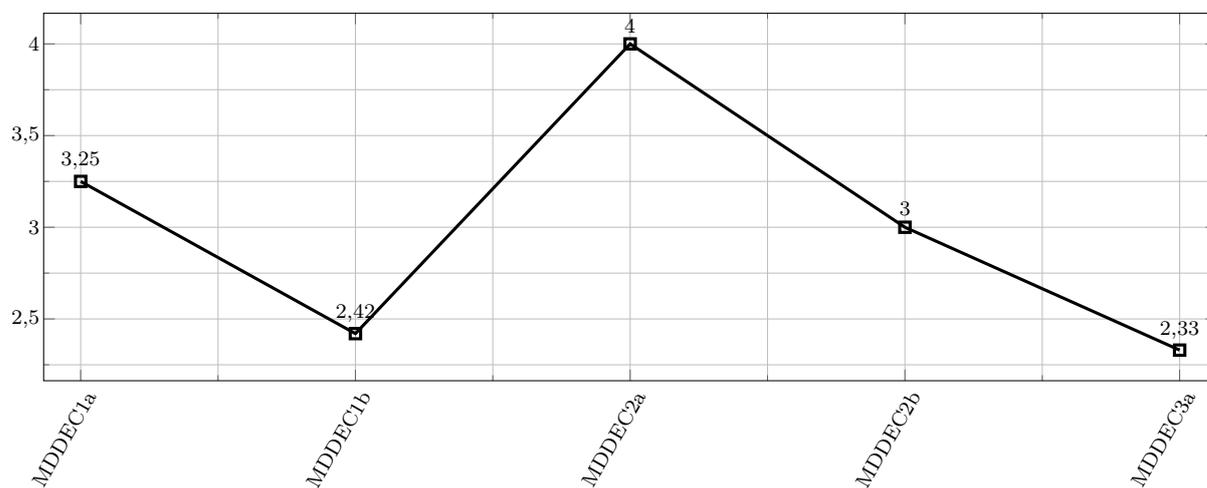
o comitê de melhor classificação ordenada por cálculo de média dos ranks e por cálculo de média dos postos. Assim, o MDDEC_{2a} apresenta-se como o melhor comitê entre os comitês avaliados. Vale ressaltar que o MDDEC_{2a} obteve média das diferenças (1,83) duas vezes maior que ao segundo colocado, o MDDEC_{1a} (0,83).

Figura 52 – Resumo das diferenças calculadas com acurácia, MTR e MCC para HT e NB. Considerando comparações com parametrizações do MDDE.

(a) Diferenças calculadas usando média das diferenças.



(b) Diferenças calculadas usando média dos postos.



Fonte: O AUTOR (2020).

Tabela 44 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,97 | 86,41 | 86,10 | 86,99 | 86,93 | 86,11 | 86,97 | 86,35 |
| A-20k-LED | 69,17 | 68,62 | 68,62 | 69,05 | 69,94 | 68,95 | 69,17 | 69,07 |
| A-20k-Mix | 90,82 | 90,82 | 90,75 | 90,84 | 90,76 | 90,44 | 90,82 | 90,75 |
| A-20k-RBF | 74,81 | 74,96 | 74,73 | 74,73 | 74,96 | 74,75 | 74,81 | 74,92 |
| A-20k-Sine | 89,77 | 89,76 | 89,72 | 89,81 | 89,75 | 89,60 | 89,77 | 89,72 |
| A-20k-Wav | 79,37 | 79,27 | 79,11 | 79,35 | 79,31 | 78,87 | 79,37 | 79,20 |
| A-130k-Agr | 91,47 | 92,03 | 91,94 | 91,34 | 92,03 | 91,91 | 91,47 | 91,94 |
| A-130k-LED | 70,61 | 71,10 | 71,17 | 70,59 | 71,55 | 71,04 | 70,61 | 71,39 |
| A-130k-Mix | 92,58 | 92,58 | 92,60 | 92,58 | 92,59 | 92,57 | 92,58 | 92,60 |
| A-130k-RBF | 77,98 | 79,63 | 79,69 | 77,84 | 79,53 | 79,63 | 77,98 | 79,63 |
| A-130k-Sine | 93,34 | 93,34 | 93,33 | 93,30 | 93,34 | 93,32 | 93,34 | 93,33 |
| A-130k-Wav | 79,77 | 79,66 | 79,89 | 79,78 | 79,70 | 79,68 | 79,77 | 79,75 |
| A-890k-Agr | 95,15 | 95,26 | 95,26 | 95,02 | 95,26 | 94,24 | 95,15 | 95,26 |
| A-890k-LED | 70,99 | 71,49 | 72,13 | 70,90 | 71,87 | 71,96 | 70,99 | 72,06 |
| A-890k-Mix | 94,95 | 94,95 | 94,97 | 94,95 | 94,97 | 94,96 | 94,95 | 94,97 |
| A-890k-RBF | 83,34 | 85,10 | 85,14 | 82,31 | 85,04 | 84,98 | 83,34 | 85,07 |
| A-890k-Sine | 96,96 | 96,96 | 96,97 | 96,92 | 96,97 | 96,97 | 96,96 | 96,97 |
| A-890k-Wav | 80,92 | 81,77 | 81,99 | 80,56 | 81,67 | 81,84 | 80,92 | 81,74 |
| G-20k-Agr | 85,22 | 84,31 | 81,30 | 85,44 | 84,96 | 82,59 | 85,22 | 83,31 |
| G-20k-LED | 68,80 | 68,58 | 68,57 | 68,76 | 69,48 | 68,61 | 68,80 | 68,81 |
| G-20k-Mix | 87,43 | 87,26 | 87,35 | 87,34 | 87,55 | 87,64 | 87,43 | 87,63 |
| G-20k-RBF | 73,91 | 73,97 | 73,58 | 73,88 | 74,10 | 73,51 | 73,91 | 73,79 |
| G-20k-Sine | 86,72 | 86,50 | 86,88 | 86,71 | 86,76 | 86,88 | 86,72 | 86,88 |
| G-20k-Wav | 78,80 | 78,54 | 78,37 | 78,87 | 78,72 | 78,33 | 78,80 | 78,55 |
| G-130k-Agr | 91,18 | 91,73 | 90,56 | 91,08 | 91,74 | 90,04 | 91,18 | 90,64 |
| G-130k-LED | 70,59 | 71,03 | 71,02 | 70,55 | 71,48 | 70,78 | 70,59 | 71,20 |
| G-130k-Mix | 92,04 | 92,03 | 92,14 | 92,01 | 92,09 | 92,15 | 92,04 | 92,14 |
| G-130k-RBF | 77,77 | 79,35 | 79,21 | 77,45 | 79,27 | 79,25 | 77,77 | 79,30 |
| G-130k-Sine | 92,79 | 92,77 | 92,71 | 92,73 | 92,81 | 92,70 | 92,79 | 92,71 |
| G-130k-Wav | 79,70 | 79,67 | 79,76 | 79,61 | 79,72 | 79,69 | 79,70 | 79,73 |
| G-890k-Agr | 95,11 | 95,23 | 93,86 | 94,98 | 95,22 | 92,30 | 95,11 | 94,26 |
| G-890k-LED | 70,99 | 71,48 | 72,06 | 70,90 | 71,87 | 71,80 | 70,99 | 72,05 |
| G-890k-Mix | 94,85 | 94,86 | 94,90 | 94,84 | 94,88 | 94,90 | 94,85 | 94,90 |
| G-890k-RBF | 83,09 | 85,07 | 84,95 | 82,50 | 84,99 | 84,89 | 83,09 | 84,95 |
| G-890k-Sine | 96,84 | 96,85 | 96,79 | 96,80 | 96,87 | 96,72 | 96,84 | 96,79 |
| G-890k-Wav | 80,85 | 81,70 | 81,94 | 80,68 | 81,68 | 81,89 | 80,85 | 81,70 |
| Airlines | 65,74 | 65,57 | 65,98 | 65,60 | 65,93 | 65,94 | 65,74 | 65,83 |
| Connect4 | 75,12 | 74,82 | 74,57 | 75,17 | 75,03 | 74,50 | 75,12 | 74,75 |
| Outdoor | 60,90 | 59,23 | 59,60 | 59,89 | 60,31 | 59,80 | 60,90 | 59,24 |
| Rialto | 38,57 | 36,19 | 27,39 | 45,06 | 43,03 | 38,85 | 38,57 | 30,52 |
| SpamData | 92,16 | 92,18 | 92,05 | 92,44 | 92,17 | 92,06 | 92,16 | 92,05 |
| WineRed | 52,13 | 53,03 | 53,02 | 52,12 | 52,71 | 53,03 | 52,13 | 53,03 |
| WineWhite | 46,00 | 44,44 | 44,73 | 46,01 | 44,41 | 45,13 | 46,00 | 44,73 |
| Vitórias | 1 | 0 | 1 | 0 | 7 | 0 | 1 | 2 |
| Derrotas | 1 | 1 | 2 | 3 | 0 | 3 | 1 | 1 |
| Média Postos | 4,57 | 4,45 | 4,48 | 5,50 | 3,13 | 5,24 | 4,57 | 4,07 |

Fonte: O AUTOR (2020).

Tabela 45 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 87,03 | 86,82 | 85,04 | 87,05 | 87,03 | 86,35 | 87,03 | 86,58 |
| A-20k-LED | 70,31 | 67,07 | 64,66 | 70,28 | 70,59 | 65,43 | 70,31 | 67,17 |
| A-20k-Mix | 91,19 | 91,19 | 91,11 | 91,23 | 91,14 | 91,03 | 91,19 | 91,11 |
| A-20k-RBF | 74,89 | 74,89 | 74,35 | 74,87 | 74,86 | 74,35 | 74,89 | 74,58 |
| A-20k-Sine | 89,19 | 89,18 | 89,14 | 89,21 | 89,16 | 89,01 | 89,19 | 89,14 |
| A-20k-Wav | 79,75 | 79,64 | 79,55 | 79,78 | 79,67 | 78,95 | 79,75 | 79,63 |
| A-130k-Agr | 88,72 | 88,91 | 88,07 | 88,71 | 88,90 | 87,52 | 88,72 | 88,67 |
| A-130k-LED | 72,17 | 71,91 | 67,29 | 72,12 | 73,05 | 67,28 | 72,17 | 71,30 |
| A-130k-Mix | 91,94 | 91,94 | 91,93 | 91,94 | 91,93 | 91,84 | 91,94 | 91,93 |
| A-130k-RBF | 75,25 | 75,30 | 74,24 | 75,26 | 75,32 | 73,04 | 75,25 | 74,79 |
| A-130k-Sine | 89,49 | 89,49 | 89,48 | 89,49 | 89,48 | 89,04 | 89,49 | 89,48 |
| A-130k-Wav | 80,33 | 80,33 | 80,30 | 80,33 | 80,32 | 80,09 | 80,33 | 80,32 |
| A-890k-Agr | 89,11 | 89,36 | 89,04 | 89,10 | 89,34 | 84,59 | 89,11 | 89,22 |
| A-890k-LED | 72,76 | 73,51 | 71,56 | 72,69 | 73,81 | 68,70 | 72,76 | 73,30 |
| A-890k-Mix | 92,08 | 92,08 | 92,07 | 92,07 | 92,07 | 89,53 | 92,08 | 92,07 |
| A-890k-RBF | 75,36 | 75,45 | 75,15 | 75,35 | 75,44 | 71,13 | 75,36 | 75,41 |
| A-890k-Sine | 89,50 | 89,50 | 89,50 | 89,50 | 89,50 | 83,66 | 89,50 | 89,50 |
| A-890k-Wav | 80,39 | 80,40 | 80,38 | 80,38 | 80,40 | 79,96 | 80,39 | 80,40 |
| G-20k-Agr | 84,05 | 81,94 | 76,51 | 84,33 | 84,12 | 79,13 | 84,05 | 80,58 |
| G-20k-LED | 69,72 | 66,22 | 62,64 | 69,91 | 69,42 | 62,78 | 69,72 | 65,80 |
| G-20k-Mix | 87,78 | 87,68 | 88,17 | 87,69 | 87,91 | 88,19 | 87,78 | 88,17 |
| G-20k-RBF | 73,82 | 73,26 | 69,56 | 73,87 | 73,76 | 70,98 | 73,82 | 71,59 |
| G-20k-Sine | 86,76 | 86,57 | 87,13 | 86,69 | 87,00 | 87,12 | 86,76 | 87,12 |
| G-20k-Wav | 78,94 | 78,93 | 78,77 | 79,02 | 78,97 | 78,12 | 78,94 | 78,78 |
| G-130k-Agr | 88,20 | 88,29 | 85,81 | 88,26 | 88,31 | 85,16 | 88,20 | 86,91 |
| G-130k-LED | 72,17 | 71,31 | 66,94 | 72,11 | 72,52 | 64,66 | 72,17 | 69,86 |
| G-130k-Mix | 91,40 | 91,39 | 91,47 | 91,40 | 91,43 | 91,05 | 91,40 | 91,47 |
| G-130k-RBF | 75,05 | 75,01 | 72,03 | 75,05 | 75,07 | 71,35 | 75,05 | 73,15 |
| G-130k-Sine | 89,20 | 89,19 | 89,29 | 89,22 | 89,31 | 87,31 | 89,20 | 89,31 |
| G-130k-Wav | 80,11 | 80,07 | 79,88 | 80,14 | 80,08 | 79,89 | 80,11 | 80,04 |
| G-890k-Agr | 89,02 | 89,26 | 86,47 | 89,01 | 89,26 | 82,34 | 89,02 | 87,56 |
| G-890k-LED | 72,75 | 73,38 | 70,45 | 72,68 | 73,53 | 65,79 | 72,75 | 72,31 |
| G-890k-Mix | 92,00 | 91,99 | 92,00 | 91,99 | 92,00 | 86,13 | 92,00 | 92,00 |
| G-890k-RBF | 75,33 | 75,32 | 74,60 | 75,32 | 75,41 | 69,84 | 75,33 | 74,93 |
| G-890k-Sine | 89,47 | 89,48 | 89,30 | 89,47 | 89,30 | 78,31 | 89,47 | 89,30 |
| G-890k-Wav | 80,35 | 80,35 | 80,17 | 80,35 | 80,34 | 79,67 | 80,35 | 80,34 |
| Airlines | 65,67 | 66,92 | 66,58 | 65,63 | 66,96 | 66,58 | 65,67 | 66,92 |
| Connect4 | 75,19 | 74,76 | 74,16 | 75,23 | 75,11 | 74,63 | 75,19 | 74,67 |
| Outdoor | 61,12 | 60,23 | 60,13 | 60,91 | 60,63 | 59,40 | 61,12 | 60,07 |
| Rialto | 38,07 | 28,98 | 25,69 | 44,55 | 34,67 | 26,61 | 38,07 | 24,29 |
| SpamData | 91,45 | 91,35 | 91,32 | 91,86 | 91,22 | 91,06 | 91,45 | 91,32 |
| WineRed | 47,87 | 47,45 | 47,70 | 47,87 | 47,51 | 47,70 | 47,87 | 47,51 |
| WineWhite | 46,73 | 45,48 | 45,03 | 46,69 | 45,62 | 42,95 | 46,73 | 45,49 |
| Vitórias | 4 | 3 | 1 | 4 | 4 | 0 | 4 | 2 |
| Derrotas | 0 | 4 | 6 | 0 | 0 | 7 | 0 | 5 |
| Média Postos | 3,27 | 4,02 | 6,23 | 3,47 | 3,28 | 7,33 | 3,27 | 5,14 |

Fonte: O AUTOR (2020).

Tabela 46 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|--------------|----------------|----------------|-----------------|----------------|--------------|--------------|---------------|
| A-20k-Agr | 122,87 | 63,53 | 42,35 | 130,68 | 104,95 | 31,53 | 122,87 | 53,47 |
| A-20k-LED | 3,15 | 1,40 | 1,77 | 3,54 | 2,51 | 1,90 | 3,15 | 2,02 |
| A-20k-Mix | 341,80 | 340,77 | 310,29 | 453,22 | 317,56 | 139,88 | 341,80 | 310,29 |
| A-20k-RBF | 31,94 | 10,12 | 4,92 | 35,02 | 16,30 | 5,29 | 31,94 | 13,36 |
| A-20k-Sine | 307,47 | 285,17 | 264,83 | 417,27 | 286,89 | 177,59 | 307,47 | 264,83 |
| A-20k-Wav | 1,00 | 0,60 | 0,51 | 1,30 | 0,80 | 0,46 | 1,00 | 0,66 |
| A-130k-Agr | 987,22 | 1058,17 | 756,50 | 1094,92 | 995,02 | 500,50 | 987,22 | 778,71 |
| A-130k-LED | 7,98 | 1,92 | 2,18 | 10,89 | 7,45 | 1,91 | 7,98 | 2,83 |
| A-130k-Mix | 1970,80 | 1970,80 | 1993,70 | 3070,98 | 2067,37 | 1413,07 | 1970,80 | 1993,70 |
| A-130k-RBF | 214,45 | 267,02 | 236,95 | 217,38 | 351,52 | 165,69 | 214,45 | 281,34 |
| A-130k-Sine | 2090,31 | 2011,33 | 1717,06 | 2977,56 | 1989,87 | 1386,42 | 2090,31 | 1717,06 |
| A-130k-Wav | 59,91 | 40,80 | 5,45 | 41,51 | 42,51 | 8,30 | 59,91 | 44,98 |
| A-890k-Agr | 8259,71 | 9114,79 | 6021,74 | 8051,83 | 6654,72 | 2128,67 | 8259,71 | 6021,74 |
| A-890k-LED | 10,30 | 1,70 | 2,14 | 13,85 | 10,79 | 3,12 | 10,30 | 3,89 |
| A-890k-Mix | 16451,68 | 16451,68 | 13538,84 | 20717,91 | 14973,11 | 10003,23 | 16451,68 | 13538,84 |
| A-890k-RBF | 1193,68 | 2130,20 | 2284,57 | 943,31 | 2092,24 | 797,49 | 1193,68 | 2090,65 |
| A-890k-Sine | 16724,42 | 16724,42 | 11830,04 | 20342,81 | 14485,30 | 8667,64 | 16724,42 | 11830,04 |
| A-890k-Wav | 129,77 | 202,51 | 89,87 | 89,67 | 159,77 | 82,92 | 129,77 | 157,78 |
| G-20k-Agr | 9,34 | 4,08 | 2,49 | 3,24 | 8,81 | 4,41 | 9,34 | 5,06 |
| G-20k-LED | 1,39 | 1,53 | 1,77 | 1,37 | 2,15 | 1,76 | 1,39 | 1,92 |
| G-20k-Mix | 5,89 | 5,07 | 1,01 | 8,06 | 0,84 | 1,04 | 5,89 | 0,93 |
| G-20k-RBF | 5,59 | 1,47 | 0,64 | 6,30 | 4,96 | 0,98 | 5,59 | 1,18 |
| G-20k-Sine | 3,77 | 2,13 | 1,54 | 3,01 | 1,05 | 6,13 | 3,77 | 1,34 |
| G-20k-Wav | 0,65 | 0,35 | 0,28 | 0,70 | 0,42 | 0,34 | 0,65 | 0,38 |
| G-130k-Agr | 93,64 | 105,74 | 113,05 | 64,79 | 180,39 | 99,82 | 93,64 | 109,76 |
| G-130k-LED | 3,04 | 1,66 | 1,78 | 2,69 | 3,20 | 1,60 | 3,04 | 1,94 |
| G-130k-Mix | 31,51 | 29,52 | 0,87 | 63,79 | 0,90 | 24,43 | 31,51 | 0,87 |
| G-130k-RBF | 47,85 | 51,99 | 31,42 | 43,01 | 60,55 | 26,93 | 47,85 | 25,67 |
| G-130k-Sine | 26,27 | 8,39 | 1,46 | 15,56 | 1,00 | 1,50 | 26,27 | 1,46 |
| G-130k-Wav | 2,62 | 1,21 | 0,40 | 3,58 | 4,09 | 0,44 | 2,62 | 0,72 |
| G-890k-Agr | 681,24 | 661,56 | 859,80 | 310,04 | 1518,90 | 547,29 | 681,24 | 887,35 |
| G-890k-LED | 2,98 | 2,67 | 1,87 | 4,15 | 7,32 | 3,25 | 2,98 | 1,73 |
| G-890k-Mix | 331,91 | 299,56 | 1,00 | 355,90 | 0,80 | 1,13 | 331,91 | 170,44 |
| G-890k-RBF | 238,84 | 550,73 | 308,92 | 158,37 | 430,97 | 201,27 | 238,84 | 375,70 |
| G-890k-Sine | 29,12 | 14,27 | 1,86 | 28,37 | 0,86 | 151,36 | 29,12 | 153,91 |
| G-890k-Wav | 16,96 | 45,73 | 1,25 | 13,04 | 59,01 | 60,30 | 16,96 | 20,62 |
| Vitórias | 2 | 3 | 1 | 3 | 3 | 0 | 2 | 2 |
| Derrotas | 0 | 0 | 6 | 0 | 0 | 7 | 0 | 3 |
| Média Postos | 3,46 | 4,53 | 6,15 | 3,25 | 3,73 | 6,39 | 3,46 | 5,04 |

Fonte: O AUTOR (2020).

Tabela 47 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|-----------------|----------------|----------------|-----------------|---------------|----------------|-----------------|---------------|
| A-20k-Agr | 134,82 | 93,90 | 60,53 | 156,82 | 118,32 | 59,89 | 134,82 | 84,73 |
| A-20k-LED | 33,70 | 0,76 | 0,37 | 35,95 | 15,77 | 0,61 | 33,70 | 0,89 |
| A-20k-Mix | 345,14 | 345,14 | 306,08 | 541,19 | 324,63 | 241,53 | 345,14 | 306,08 |
| A-20k-RBF | 23,48 | 9,94 | 3,17 | 28,72 | 14,65 | 6,02 | 23,48 | 10,93 |
| A-20k-Sine | 317,21 | 293,51 | 261,41 | 387,25 | 288,20 | 174,08 | 317,21 | 265,86 |
| A-20k-Wav | 52,83 | 29,56 | 23,88 | 69,24 | 40,73 | 7,49 | 52,83 | 31,43 |
| A-130k-Agr | 699,48 | 718,22 | 471,37 | 693,56 | 673,58 | 261,16 | 699,48 | 587,64 |
| A-130k-LED | 31,04 | 21,75 | 1,45 | 31,45 | 100,88 | 9,34 | 31,04 | 36,61 |
| A-130k-Mix | 2277,70 | 2277,70 | 2037,77 | 2572,16 | 2099,09 | 911,35 | 2277,70 | 2037,77 |
| A-130k-RBF | 82,63 | 137,09 | 28,68 | 112,29 | 166,44 | 5,14 | 82,63 | 129,44 |
| A-130k-Sine | 2154,98 | 1987,05 | 1740,78 | 2290,29 | 1811,29 | 501,97 | 2154,98 | 1768,44 |
| A-130k-Wav | 465,99 | 454,95 | 368,31 | 466,26 | 463,13 | 148,40 | 465,99 | 467,19 |
| A-890k-Agr | 4211,14 | 4598,58 | 3407,33 | 4119,36 | 4024,34 | 412,88 | 4211,14 | 3919,18 |
| A-890k-LED | 44,24 | 183,16 | 64,25 | 45,04 | 365,71 | 63,20 | 44,24 | 226,65 |
| A-890k-Mix | 14959,61 | 14959,61 | 13633,88 | 15699,45 | 13760,01 | 3255,51 | 14959,61 | 13633,88 |
| A-890k-RBF | 232,20 | 422,33 | 202,49 | 265,87 | 415,04 | 1,82 | 232,20 | 465,75 |
| A-890k-Sine | 13353,35 | 12703,57 | 11314,39 | 13089,69 | 11429,14 | 1149,66 | 13353,35 | 11424,12 |
| A-890k-Wav | 1122,11 | 1680,90 | 2061,09 | 844,94 | 1841,17 | 650,28 | 1122,11 | 1866,92 |
| G-20k-Agr | 8,47 | 4,44 | 5,81 | 9,19 | 10,37 | 10,87 | 8,47 | 11,07 |
| G-20k-LED | 4,16 | 0,52 | 0,16 | 5,85 | 2,82 | 0,20 | 4,16 | 0,52 |
| G-20k-Mix | 4,39 | 3,18 | 1,06 | 5,06 | 0,79 | 1,06 | 4,39 | 1,06 |
| G-20k-RBF | 5,30 | 2,80 | 1,12 | 5,74 | 7,34 | 1,62 | 5,30 | 2,12 |
| G-20k-Sine | 2,42 | 2,33 | 6,92 | 2,16 | 0,89 | 4,08 | 2,42 | 3,96 |
| G-20k-Wav | 0,44 | 0,30 | 0,30 | 0,52 | 0,35 | 0,26 | 0,44 | 0,30 |
| G-130k-Agr | 41,71 | 66,46 | 93,86 | 52,96 | 142,62 | 99,09 | 41,71 | 104,57 |
| G-130k-LED | 8,12 | 5,35 | 0,84 | 8,86 | 33,65 | 0,37 | 8,12 | 4,12 |
| G-130k-Mix | 33,48 | 29,66 | 0,86 | 58,45 | 0,93 | 89,50 | 33,48 | 0,88 |
| G-130k-RBF | 16,89 | 19,68 | 4,31 | 19,05 | 22,57 | 8,92 | 16,89 | 14,58 |
| G-130k-Sine | 7,23 | 9,97 | 1,41 | 8,36 | 0,94 | 59,55 | 7,23 | 1,10 |
| G-130k-Wav | 78,12 | 59,29 | 40,27 | 69,85 | 66,45 | 49,33 | 78,12 | 62,78 |
| G-890k-Agr | 277,70 | 500,11 | 444,76 | 357,04 | 807,85 | 185,57 | 277,70 | 487,40 |
| G-890k-LED | 10,25 | 41,38 | 2,25 | 14,84 | 111,74 | 0,56 | 10,25 | 13,99 |
| G-890k-Mix | 235,23 | 221,04 | 0,78 | 272,03 | 0,80 | 1239,11 | 235,23 | 0,81 |
| G-890k-RBF | 19,41 | 102,06 | 33,19 | 19,49 | 118,59 | 12,43 | 19,41 | 71,12 |
| G-890k-Sine | 11,52 | 14,18 | 137,11 | 19,65 | 1,14 | 271,11 | 11,52 | 1,21 |
| G-890k-Wav | 106,85 | 246,68 | 373,71 | 47,64 | 226,62 | 343,98 | 106,85 | 250,62 |
| Vitórias | 2 | 3 | 1 | 4 | 3 | 0 | 2 | 2 |
| Derrotas | 1 | 0 | 6 | 0 | 0 | 7 | 1 | 2 |
| Média Postos | 3,96 | 4,10 | 6,15 | 2,97 | 3,94 | 6,25 | 3,96 | 4,67 |

Fonte: O AUTOR (2020).

Tabela 48 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| A-20k-Agr | 0,87 | 0,84 | 0,81 | 0,87 | 0,94 | 0,73 | 0,87 | 0,86 |
| A-20k-LED | 0,17 | 0,04 | 0,01 | 0,17 | 0,14 | 0,05 | 0,17 | 0,07 |
| A-20k-Mix | 0,95 | 0,95 | 1,00 | 0,94 | 1,00 | 1,00 | 0,95 | 1,00 |
| A-20k-RBF | 0,64 | 0,66 | 0,62 | 0,60 | 0,66 | 0,53 | 0,64 | 0,68 |
| A-20k-Sine | 0,98 | 0,98 | 1,00 | 0,96 | 1,00 | 1,00 | 0,98 | 1,00 |
| A-20k-Wav | 0,32 | 0,35 | 0,37 | 0,33 | 0,38 | 0,19 | 0,32 | 0,34 |
| A-130k-Agr | 0,82 | 0,94 | 0,90 | 0,80 | 0,97 | 0,83 | 0,82 | 0,90 |
| A-130k-LED | 0,08 | 0,02 | 0,01 | 0,09 | 0,06 | 0,01 | 0,08 | 0,03 |
| A-130k-Mix | 0,94 | 0,94 | 1,00 | 0,93 | 0,98 | 0,99 | 0,94 | 1,00 |
| A-130k-RBF | 0,45 | 0,70 | 0,72 | 0,44 | 0,73 | 0,63 | 0,45 | 0,73 |
| A-130k-Sine | 0,98 | 0,98 | 1,00 | 0,94 | 1,00 | 0,99 | 0,98 | 1,00 |
| A-130k-Wav | 0,28 | 0,35 | 0,43 | 0,24 | 0,34 | 0,39 | 0,28 | 0,36 |
| A-890k-Agr | 0,84 | 0,98 | 0,95 | 0,77 | 0,98 | 0,69 | 0,84 | 0,95 |
| A-890k-LED | 0,03 | 0,01 | 0,02 | 0,04 | 0,03 | 0,01 | 0,03 | 0,01 |
| A-890k-Mix | 0,94 | 0,94 | 1,00 | 0,93 | 0,98 | 0,98 | 0,94 | 1,00 |
| A-890k-RBF | 0,33 | 0,72 | 0,76 | 0,26 | 0,70 | 0,50 | 0,33 | 0,74 |
| A-890k-Sine | 0,98 | 0,98 | 1,00 | 0,93 | 1,00 | 0,99 | 0,98 | 1,00 |
| A-890k-Wav | 0,17 | 0,26 | 0,35 | 0,12 | 0,22 | 0,19 | 0,17 | 0,24 |
| Vitórias | 1 | 1 | 4 | 0 | 5 | 0 | 1 | 5 |
| Derrotas | 3 | 2 | 0 | 6 | 0 | 3 | 3 | 0 |
| Média Postos | 5,31 | 4,67 | 3,37 | 6,33 | 2,72 | 5,22 | 5,31 | 3,08 |

Fonte: O AUTOR (2020).

Tabela 49 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|--------------|--------|-------------|-------------|-------------|-------------|-------------|--------|-------------|
| A-20k-Agr | 0,90 | 0,87 | 0,79 | 0,89 | 0,93 | 0,84 | 0,90 | 0,89 |
| A-20k-LED | 0,37 | 0,13 | 0,07 | 0,38 | 0,27 | 0,10 | 0,37 | 0,16 |
| A-20k-Mix | 0,99 | 0,99 | 1,00 | 0,99 | 1,00 | 1,00 | 0,99 | 1,00 |
| A-20k-RBF | 0,66 | 0,73 | 0,71 | 0,61 | 0,73 | 0,61 | 0,66 | 0,74 |
| A-20k-Sine | 0,99 | 0,99 | 1,00 | 0,99 | 1,00 | 1,00 | 0,99 | 1,00 |
| A-20k-Wav | 0,70 | 0,62 | 0,59 | 0,75 | 0,66 | 0,37 | 0,70 | 0,64 |
| A-130k-Agr | 0,62 | 0,90 | 0,83 | 0,60 | 0,89 | 0,61 | 0,62 | 0,90 |
| A-130k-LED | 0,16 | 0,19 | 0,12 | 0,16 | 0,26 | 0,06 | 0,16 | 0,24 |
| A-130k-Mix | 0,99 | 0,99 | 1,00 | 0,99 | 1,00 | 0,95 | 0,99 | 1,00 |
| A-130k-RBF | 0,33 | 0,60 | 0,72 | 0,33 | 0,59 | 0,33 | 0,33 | 0,64 |
| A-130k-Sine | 0,98 | 0,98 | 1,00 | 0,96 | 0,99 | 0,73 | 0,98 | 0,99 |
| A-130k-Wav | 0,67 | 0,68 | 0,65 | 0,64 | 0,69 | 0,38 | 0,67 | 0,71 |
| A-890k-Agr | 0,28 | 0,83 | 0,84 | 0,28 | 0,79 | 0,41 | 0,28 | 0,84 |
| A-890k-LED | 0,07 | 0,12 | 0,07 | 0,08 | 0,17 | 0,08 | 0,07 | 0,15 |
| A-890k-Mix | 0,98 | 0,98 | 1,00 | 0,97 | 1,00 | 0,61 | 0,98 | 1,00 |
| A-890k-RBF | 0,14 | 0,38 | 0,55 | 0,14 | 0,37 | 0,18 | 0,14 | 0,40 |
| A-890k-Sine | 0,89 | 0,92 | 0,97 | 0,85 | 0,95 | 0,47 | 0,89 | 0,96 |
| A-890k-Wav | 0,33 | 0,46 | 0,58 | 0,31 | 0,46 | 0,25 | 0,33 | 0,47 |
| Vitórias | 2 | 4 | 1 | 1 | 5 | 0 | 2 | 5 |
| Derrotas | 3 | 2 | 0 | 5 | 0 | 7 | 3 | 0 |
| Média Postos | 55,28 | 4,22 | 3,70 | 5,86 | 2,70 | 6,59 | 5,28 | 2,39 |

Fonte: O AUTOR (2020).

Tabela 50 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 85,89 | 87,00 | 86,14 | 86,98 | 86,92 | 86,16 | 86,97 | 86,36 |
| A-20k-LED | 40,90 | 69,53 | 68,73 | 68,99 | 70,09 | 69,06 | 69,17 | 69,11 |
| A-20k-Mix | 90,21 | 90,82 | 90,76 | 90,83 | 90,78 | 90,53 | 90,82 | 90,76 |
| A-20k-RBF | 73,90 | 74,77 | 74,75 | 74,71 | 74,90 | 74,77 | 74,81 | 74,87 |
| A-20k-Sine | 89,30 | 89,77 | 89,73 | 89,75 | 89,76 | 89,62 | 89,77 | 89,73 |
| A-20k-Wav | 76,90 | 79,37 | 79,12 | 79,33 | 79,29 | 78,90 | 79,37 | 79,18 |
| A-130k-Agr | 87,68 | 91,61 | 91,94 | 91,22 | 92,00 | 91,93 | 91,47 | 91,95 |
| A-130k-LED | 41,56 | 71,11 | 71,23 | 70,60 | 71,72 | 71,21 | 70,61 | 71,46 |
| A-130k-Mix | 91,26 | 92,58 | 92,61 | 92,55 | 92,60 | 92,58 | 92,58 | 92,61 |
| A-130k-RBF | 74,30 | 77,95 | 79,66 | 77,57 | 79,34 | 79,64 | 77,98 | 79,62 |
| A-130k-Sine | 91,50 | 93,34 | 93,34 | 93,25 | 93,33 | 93,33 | 93,34 | 93,34 |
| A-130k-Wav | 77,49 | 79,72 | 79,90 | 79,75 | 79,76 | 79,72 | 79,77 | 79,74 |
| A-890k-Agr | 88,96 | 95,18 | 95,26 | 94,95 | 95,26 | 94,28 | 95,15 | 95,26 |
| A-890k-LED | 41,53 | 71,39 | 72,13 | 70,90 | 71,97 | 72,01 | 70,99 | 72,09 |
| A-890k-Mix | 92,11 | 94,95 | 94,97 | 94,93 | 94,97 | 94,97 | 94,95 | 94,97 |
| A-890k-RBF | 74,67 | 83,58 | 85,12 | 81,70 | 84,97 | 85,01 | 83,34 | 85,08 |
| A-890k-Sine | 94,29 | 96,96 | 96,98 | 96,90 | 96,99 | 96,96 | 96,96 | 96,98 |
| A-890k-Wav | 77,59 | 81,01 | 81,99 | 80,32 | 81,62 | 81,91 | 80,92 | 81,74 |
| G-20k-Agr | 84,42 | 85,22 | 81,25 | 85,44 | 85,07 | 82,63 | 85,22 | 83,23 |
| G-20k-LED | 40,86 | 69,21 | 68,58 | 68,73 | 69,66 | 68,67 | 68,80 | 68,89 |
| G-20k-Mix | 86,92 | 87,42 | 87,23 | 87,38 | 87,41 | 87,32 | 87,43 | 87,33 |
| G-20k-RBF | 73,04 | 74,01 | 73,57 | 73,84 | 74,03 | 73,62 | 73,91 | 73,77 |
| G-20k-Sine | 86,10 | 86,65 | 86,65 | 86,69 | 86,72 | 86,60 | 86,72 | 86,54 |
| G-20k-Wav | 76,55 | 78,76 | 78,37 | 78,84 | 78,77 | 78,29 | 78,80 | 78,55 |
| G-130k-Agr | 87,48 | 91,31 | 90,60 | 90,96 | 91,71 | 89,74 | 91,18 | 90,82 |
| G-130k-LED | 41,60 | 71,07 | 70,78 | 70,54 | 71,65 | 70,57 | 70,59 | 71,24 |
| G-130k-Mix | 90,64 | 92,04 | 92,11 | 91,97 | 92,08 | 92,10 | 92,04 | 92,11 |
| G-130k-RBF | 74,19 | 77,69 | 79,22 | 77,28 | 79,19 | 79,31 | 77,77 | 79,31 |
| G-130k-Sine | 91,03 | 92,78 | 92,67 | 92,69 | 92,80 | 92,82 | 92,79 | 92,67 |
| G-130k-Wav | 77,47 | 79,69 | 79,76 | 79,66 | 79,75 | 79,77 | 79,70 | 79,74 |
| G-890k-Agr | 88,93 | 95,14 | 93,36 | 94,90 | 95,22 | 92,10 | 95,11 | 93,70 |
| G-890k-LED | 41,54 | 71,37 | 72,02 | 70,89 | 71,96 | 71,70 | 70,99 | 72,06 |
| G-890k-Mix | 91,99 | 94,86 | 94,88 | 94,81 | 94,88 | 94,89 | 94,85 | 94,88 |
| G-890k-RBF | 74,65 | 83,44 | 84,96 | 81,82 | 84,81 | 84,90 | 83,09 | 85,03 |
| G-890k-Sine | 94,12 | 96,84 | 96,74 | 96,79 | 96,89 | 96,73 | 96,84 | 96,74 |
| G-890k-Wav | 77,58 | 81,02 | 81,93 | 80,30 | 81,57 | 81,82 | 80,85 | 81,70 |
| Airlines | 64,30 | 66,02 | 66,04 | 65,60 | 65,98 | 66,30 | 65,74 | 66,01 |
| Connect4 | 74,73 | 75,14 | 74,47 | 75,15 | 75,09 | 74,66 | 75,12 | 74,88 |
| Outdoor | 57,74 | 61,53 | 58,79 | 59,93 | 60,18 | 58,96 | 60,90 | 59,14 |
| Rialto | 41,97 | 38,55 | 31,04 | 42,83 | 38,65 | 33,41 | 38,57 | 34,45 |
| SpamData | 92,00 | 92,16 | 91,77 | 92,26 | 92,19 | 91,86 | 92,16 | 91,77 |
| WineRed | 52,09 | 52,56 | 52,93 | 52,22 | 52,55 | 53,03 | 52,13 | 53,03 |
| WineWhite | 47,17 | 46,49 | 44,63 | 46,85 | 44,42 | 45,13 | 46,00 | 44,42 |
| Vitórias | 0 | 3 | 1 | 1 | 7 | 1 | 2 | 3 |
| Derrotas | 7 | 1 | 2 | 3 | 0 | 2 | 2 | 1 |
| Média Postos | 7,51 | 3,77 | 4,24 | 5,02 | 2,93 | 4,67 | 4,05 | 3,80 |

Fonte: O AUTOR (2020).

Tabela 51 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,05 | 87,05 | 84,97 | 87,06 | 87,02 | 86,40 | 87,03 | 86,58 |
| A-20k-LED | 43,06 | 70,76 | 64,69 | 70,28 | 70,77 | 65,54 | 70,31 | 67,36 |
| A-20k-Mix | 90,76 | 91,19 | 91,11 | 91,23 | 91,15 | 91,06 | 91,19 | 91,11 |
| A-20k-RBF | 74,07 | 74,90 | 74,35 | 74,86 | 74,81 | 74,37 | 74,89 | 74,56 |
| A-20k-Sine | 88,94 | 89,19 | 89,15 | 89,21 | 89,17 | 89,02 | 89,19 | 89,15 |
| A-20k-Wav | 77,70 | 79,70 | 79,55 | 79,78 | 79,65 | 78,88 | 79,75 | 79,62 |
| A-130k-Agr | 87,71 | 88,76 | 88,07 | 88,71 | 88,89 | 87,52 | 88,72 | 88,68 |
| A-130k-LED | 42,71 | 72,60 | 67,13 | 72,05 | 73,25 | 67,29 | 72,17 | 71,21 |
| A-130k-Mix | 91,70 | 91,94 | 91,93 | 91,94 | 91,93 | 91,85 | 91,94 | 91,93 |
| A-130k-RBF | 74,37 | 75,27 | 74,24 | 75,24 | 75,28 | 73,17 | 75,25 | 74,76 |
| A-130k-Sine | 89,23 | 89,49 | 89,48 | 89,49 | 89,49 | 89,04 | 89,49 | 89,48 |
| A-130k-Wav | 78,42 | 80,33 | 80,31 | 80,31 | 80,32 | 80,12 | 80,33 | 80,32 |
| A-890k-Agr | 88,05 | 89,14 | 89,04 | 89,09 | 89,33 | 84,23 | 89,11 | 89,23 |
| A-890k-LED | 43,08 | 73,05 | 71,56 | 72,65 | 73,80 | 68,66 | 72,76 | 73,31 |
| A-890k-Mix | 91,84 | 92,08 | 92,07 | 92,07 | 92,07 | 89,53 | 92,08 | 92,07 |
| A-890k-RBF | 74,42 | 75,37 | 75,15 | 75,34 | 75,45 | 70,98 | 75,36 | 75,41 |
| A-890k-Sine | 89,32 | 89,50 | 89,50 | 89,50 | 89,50 | 82,90 | 89,50 | 89,50 |
| A-890k-Wav | 78,44 | 80,39 | 80,38 | 80,37 | 80,39 | 79,99 | 80,39 | 80,40 |
| G-20k-Agr | 83,83 | 83,88 | 76,23 | 84,70 | 84,72 | 78,08 | 84,05 | 80,11 |
| G-20k-LED | 42,94 | 70,11 | 62,64 | 69,87 | 69,82 | 62,89 | 69,72 | 65,96 |
| G-20k-Mix | 87,19 | 87,78 | 87,89 | 87,71 | 87,76 | 87,95 | 87,78 | 87,89 |
| G-20k-RBF | 73,20 | 73,81 | 69,48 | 73,86 | 73,61 | 70,77 | 73,82 | 71,65 |
| G-20k-Sine | 85,99 | 86,75 | 86,63 | 86,75 | 86,91 | 86,69 | 86,76 | 86,60 |
| G-20k-Wav | 77,42 | 78,95 | 78,77 | 79,05 | 78,97 | 78,11 | 78,94 | 78,76 |
| G-130k-Agr | 87,32 | 88,23 | 85,67 | 88,26 | 88,33 | 85,06 | 88,20 | 86,74 |
| G-130k-LED | 42,69 | 72,51 | 66,89 | 72,05 | 72,77 | 64,76 | 72,17 | 70,26 |
| G-130k-Mix | 91,14 | 91,40 | 91,42 | 91,40 | 91,42 | 91,02 | 91,40 | 91,42 |
| G-130k-RBF | 74,22 | 75,07 | 72,09 | 75,05 | 75,05 | 71,58 | 75,05 | 73,21 |
| G-130k-Sine | 88,78 | 89,20 | 89,30 | 89,21 | 89,31 | 87,00 | 89,20 | 89,31 |
| G-130k-Wav | 78,40 | 80,10 | 79,86 | 80,13 | 80,08 | 79,87 | 80,11 | 80,03 |
| G-890k-Agr | 87,98 | 89,05 | 86,46 | 89,00 | 89,22 | 82,36 | 89,02 | 87,55 |
| G-890k-LED | 43,07 | 73,04 | 70,52 | 72,64 | 73,73 | 65,79 | 72,75 | 72,19 |
| G-890k-Mix | 91,75 | 92,00 | 92,00 | 91,99 | 92,00 | 86,13 | 92,00 | 92,00 |
| G-890k-RBF | 74,40 | 75,35 | 74,60 | 75,31 | 75,42 | 70,06 | 75,33 | 74,93 |
| G-890k-Sine | 89,25 | 89,48 | 89,31 | 89,47 | 89,50 | 77,64 | 89,47 | 89,31 |
| G-890k-Wav | 78,43 | 80,35 | 80,17 | 80,34 | 80,34 | 79,67 | 80,35 | 80,33 |
| Airlines | 63,43 | 66,25 | 66,62 | 65,65 | 66,97 | 66,53 | 65,67 | 66,93 |
| Connect4 | 74,66 | 75,17 | 74,03 | 75,17 | 75,02 | 74,62 | 75,19 | 74,46 |
| Outdoor | 59,58 | 60,59 | 59,16 | 60,94 | 60,32 | 59,78 | 61,12 | 60,78 |
| Rialto | 41,97 | 31,43 | 24,91 | 42,81 | 30,13 | 25,06 | 38,07 | 24,53 |
| SpamData | 91,31 | 91,45 | 91,32 | 91,76 | 91,58 | 91,02 | 91,45 | 91,32 |
| WineRed | 47,73 | 47,51 | 47,70 | 47,90 | 47,35 | 47,70 | 47,87 | 47,56 |
| WineWhite | 47,77 | 47,27 | 45,29 | 47,37 | 45,90 | 42,95 | 46,73 | 45,97 |
| Vitórias | 0 | 5 | 2 | 4 | 4 | 0 | 4 | 3 |
| Derrotas | 6 | 0 | 5 | 0 | 0 | 6 | 1 | 4 |
| Média Postos | 6,79 | 2,84 | 5,84 | 3,23 | 2,79 | 6,87 | 3,01 | 4,63 |

Fonte: O AUTOR (2020).

Tabela 52 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|--------------|----------------|----------------|-----------------|----------------|--------------|--------------|---------------|
| A-20k-Agr | 81,45 | 127,60 | 42,50 | 136,28 | 109,26 | 33,26 | 122,87 | 53,70 |
| A-20k-LED | 1,74 | 2,63 | 1,80 | 3,94 | 3,35 | 1,93 | 3,15 | 1,96 |
| A-20k-Mix | 205,44 | 341,80 | 310,62 | 460,72 | 317,72 | 160,99 | 341,80 | 310,62 |
| A-20k-RBF | 15,33 | 33,80 | 4,97 | 37,61 | 16,88 | 5,37 | 31,94 | 12,23 |
| A-20k-Sine | 220,08 | 307,47 | 265,28 | 418,65 | 287,41 | 177,38 | 307,47 | 265,28 |
| A-20k-Wav | 7,75 | 1,01 | 0,52 | 3,72 | 0,84 | 0,49 | 1,00 | 0,68 |
| A-130k-Agr | 422,88 | 988,36 | 753,65 | 1124,27 | 1003,12 | 480,81 | 987,22 | 776,95 |
| A-130k-LED | 3,48 | 8,57 | 2,12 | 12,51 | 10,73 | 2,34 | 7,98 | 2,80 |
| A-130k-Mix | 373,56 | 1970,80 | 2009,82 | 2750,38 | 2081,44 | 1389,62 | 1970,80 | 2009,82 |
| A-130k-RBF | 18,89 | 212,12 | 233,81 | 197,42 | 315,34 | 163,26 | 214,45 | 272,32 |
| A-130k-Sine | 542,22 | 2090,31 | 1715,36 | 2843,28 | 1992,19 | 1414,39 | 2090,31 | 1715,36 |
| A-130k-Wav | 7,71 | 35,80 | 1,33 | 40,24 | 41,07 | 4,33 | 59,91 | 42,16 |
| A-890k-Agr | 3024,32 | 8441,33 | 5999,86 | 8070,49 | 6568,56 | 2082,49 | 8259,71 | 5998,76 |
| A-890k-LED | 1,65 | 13,08 | 1,72 | 12,41 | 11,53 | 1,63 | 10,30 | 4,93 |
| A-890k-Mix | 515,27 | 16451,68 | 13550,19 | 19455,01 | 15188,82 | 9897,96 | 16451,68 | 13550,19 |
| A-890k-RBF | 21,97 | 1334,61 | 2274,43 | 729,11 | 1798,96 | 915,66 | 1193,68 | 2050,98 |
| A-890k-Sine | 1497,13 | 16724,42 | 11844,72 | 19344,26 | 14406,21 | 8808,42 | 16724,42 | 11844,72 |
| A-890k-Wav | 10,86 | 196,23 | 121,04 | 81,37 | 173,85 | 148,02 | 129,77 | 98,82 |
| G-20k-Agr | 8,69 | 9,87 | 2,43 | 2,01 | 10,11 | 4,34 | 9,34 | 5,15 |
| G-20k-LED | 1,46 | 1,72 | 1,77 | 1,48 | 2,17 | 1,75 | 1,39 | 1,90 |
| G-20k-Mix | 7,07 | 5,17 | 0,94 | 7,84 | 0,85 | 3,47 | 5,89 | 1,01 |
| G-20k-RBF | 3,11 | 5,67 | 0,70 | 5,30 | 4,76 | 1,18 | 5,59 | 1,15 |
| G-20k-Sine | 8,77 | 3,44 | 4,06 | 4,64 | 1,08 | 1,38 | 3,77 | 1,38 |
| G-20k-Wav | 1,28 | 0,65 | 0,26 | 0,89 | 0,52 | 0,27 | 0,65 | 0,36 |
| G-130k-Agr | 40,96 | 119,00 | 104,98 | 31,20 | 171,53 | 89,39 | 93,64 | 112,94 |
| G-130k-LED | 1,23 | 4,48 | 1,49 | 5,29 | 3,99 | 1,40 | 3,04 | 2,47 |
| G-130k-Mix | 22,21 | 28,96 | 0,81 | 68,73 | 0,98 | 27,30 | 31,51 | 0,81 |
| G-130k-RBF | 2,67 | 45,65 | 26,71 | 35,32 | 52,50 | 30,50 | 47,85 | 39,22 |
| G-130k-Sine | 26,68 | 13,91 | 1,46 | 16,11 | 1,11 | 1,48 | 26,27 | 1,45 |
| G-130k-Wav | 0,60 | 2,50 | 0,39 | 2,53 | 3,63 | 0,43 | 2,62 | 0,77 |
| G-890k-Agr | 499,23 | 696,99 | 663,41 | 279,04 | 1699,71 | 432,25 | 681,24 | 699,65 |
| G-890k-LED | 0,93 | 4,05 | 1,72 | 4,41 | 4,16 | 1,06 | 2,98 | 3,37 |
| G-890k-Mix | 28,03 | 334,12 | 220,65 | 348,87 | 0,93 | 220,72 | 331,91 | 220,65 |
| G-890k-RBF | 3,39 | 243,12 | 345,07 | 121,24 | 383,51 | 314,94 | 238,84 | 417,39 |
| G-890k-Sine | 123,43 | 27,85 | 324,07 | 50,25 | 0,92 | 1,64 | 29,12 | 324,07 |
| G-890k-Wav | 1,21 | 18,84 | 6,97 | 9,15 | 32,86 | 59,23 | 16,96 | 24,26 |
| Vitórias | 0 | 4 | 2 | 3 | 4 | 0 | 4 | 3 |
| Derrotas | 6 | 0 | 5 | 0 | 0 | 6 | 0 | 3 |
| Média Postos | 6,03 | 3,15 | 5,68 | 3,11 | 3,53 | 6,26 | 3,54 | 4,69 |

Fonte: O AUTOR (2020).

Tabela 53 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|-------------|-----------------|----------------|----------------|---------------|----------------|-----------------|---------------|
| A-20k-Agr | 87,13 | 134,43 | 60,54 | 160,43 | 125,47 | 59,90 | 134,82 | 84,66 |
| A-20k-LED | 2,55 | 32,14 | 0,39 | 35,13 | 16,00 | 0,66 | 33,70 | 0,95 |
| A-20k-Mix | 267,70 | 345,14 | 306,08 | 548,94 | 325,55 | 240,50 | 345,14 | 306,08 |
| A-20k-RBF | 11,84 | 24,86 | 3,17 | 31,00 | 15,17 | 5,85 | 23,48 | 10,93 |
| A-20k-Sine | 223,59 | 317,21 | 261,53 | 379,39 | 288,26 | 165,61 | 317,21 | 265,79 |
| A-20k-Wav | 13,89 | 45,71 | 23,88 | 74,46 | 43,38 | 6,61 | 52,83 | 31,38 |
| A-130k-Agr | 370,53 | 716,36 | 471,37 | 711,45 | 675,38 | 252,37 | 699,48 | 582,26 |
| A-130k-LED | 3,63 | 44,43 | 1,41 | 29,87 | 101,39 | 14,28 | 31,04 | 38,56 |
| A-130k-Mix | 1210,91 | 2277,70 | 2037,77 | 2602,19 | 2102,18 | 910,89 | 2277,70 | 2037,77 |
| A-130k-RBF | 15,33 | 93,04 | 28,68 | 111,67 | 151,33 | 13,91 | 82,63 | 128,73 |
| A-130k-Sine | 542,49 | 2154,10 | 1741,35 | 2197,27 | 1814,46 | 481,99 | 2154,98 | 1768,44 |
| A-130k-Wav | 10,58 | 417,25 | 368,36 | 446,46 | 455,83 | 152,14 | 465,99 | 463,16 |
| A-890k-Agr | 2268,08 | 4433,59 | 3407,36 | 4080,56 | 3886,66 | 382,64 | 4211,14 | 3924,75 |
| A-890k-LED | 2,34 | 59,47 | 64,27 | 44,40 | 302,68 | 49,06 | 44,24 | 185,36 |
| A-890k-Mix | 1121,19 | 14959,61 | 13633,88 | 14890,12 | 13760,01 | 3255,37 | 14959,61 | 13633,88 |
| A-890k-RBF | 16,95 | 219,60 | 202,49 | 227,77 | 418,73 | 1,84 | 232,20 | 462,39 |
| A-890k-Sine | 609,80 | 13242,88 | 11314,39 | 12057,97 | 11344,26 | 1099,07 | 13353,35 | 11424,12 |
| A-890k-Wav | 16,62 | 911,74 | 2061,09 | 779,85 | 1723,75 | 650,20 | 1122,11 | 1853,47 |
| G-20k-Agr | 6,25 | 7,28 | 5,78 | 7,46 | 13,50 | 9,81 | 8,47 | 11,32 |
| G-20k-LED | 1,18 | 4,93 | 0,16 | 5,30 | 4,83 | 0,21 | 4,16 | 0,55 |
| G-20k-Mix | 8,56 | 3,91 | 1,07 | 5,90 | 0,83 | 1,07 | 4,39 | 1,07 |
| G-20k-RBF | 3,43 | 5,33 | 1,17 | 4,31 | 7,06 | 1,44 | 5,30 | 2,29 |
| G-20k-Sine | 9,09 | 2,25 | 4,03 | 2,38 | 0,98 | 4,11 | 2,42 | 3,94 |
| G-20k-Wav | 0,96 | 0,43 | 0,30 | 0,74 | 0,36 | 0,26 | 0,44 | 0,30 |
| G-130k-Agr | 15,97 | 54,16 | 90,03 | 41,24 | 147,32 | 100,31 | 41,71 | 100,70 |
| G-130k-LED | 1,37 | 10,49 | 0,82 | 8,54 | 29,13 | 0,39 | 8,12 | 4,69 |
| G-130k-Mix | 35,16 | 29,43 | 0,87 | 65,84 | 1,09 | 86,89 | 33,48 | 0,87 |
| G-130k-RBF | 2,06 | 16,86 | 4,36 | 17,61 | 26,24 | 2,87 | 16,89 | 16,61 |
| G-130k-Sine | 23,08 | 12,62 | 1,30 | 16,18 | 1,06 | 70,63 | 7,23 | 1,18 |
| G-130k-Wav | 1,17 | 59,76 | 40,20 | 65,02 | 63,07 | 51,63 | 78,12 | 62,15 |
| G-890k-Agr | 82,38 | 317,29 | 461,68 | 282,77 | 805,55 | 213,44 | 277,70 | 506,21 |
| G-890k-LED | 1,22 | 15,46 | 2,48 | 13,71 | 93,01 | 0,56 | 10,25 | 21,62 |
| G-890k-Mix | 58,88 | 233,41 | 0,80 | 274,50 | 0,80 | 1241,08 | 235,23 | 0,81 |
| G-890k-RBF | 3,14 | 30,49 | 31,68 | 17,98 | 131,18 | 12,51 | 19,41 | 70,93 |
| G-890k-Sine | 34,61 | 11,23 | 155,69 | 35,89 | 1,44 | 145,47 | 11,52 | 55,44 |
| G-890k-Wav | 1,88 | 113,82 | 373,71 | 39,29 | 229,37 | 345,80 | 106,85 | 255,98 |
| Vitórias | 0 | 3 | 2 | 3 | 4 | 0 | 3 | 3 |
| Derrotas | 6 | 0 | 5 | 0 | 0 | 6 | 0 | 1 |
| Média Postos | 6,14 | 3,58 | 5,67 | 3,17 | 3,60 | 6,08 | 3,50 | 4,26 |

Fonte: O AUTOR (2020).

Tabela 54 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|----------|-------------|-------------|-------------|-------------|-------------|----------|-------------|
| A-20k-Agr | 0,41 | 0,88 | 0,82 | 0,86 | 0,93 | 0,76 | 0,87 | 0,86 |
| A-20k-LED | 0,08 | 0,16 | 0,01 | 0,18 | 0,16 | 0,06 | 0,17 | 0,06 |
| A-20k-Mix | 0,49 | 0,95 | 1,00 | 0,92 | 1,00 | 1,00 | 0,95 | 1,00 |
| A-20k-RBF | 0,30 | 0,63 | 0,62 | 0,60 | 0,66 | 0,54 | 0,64 | 0,68 |
| A-20k-Sine | 0,58 | 0,98 | 1,00 | 0,93 | 1,00 | 1,00 | 0,98 | 1,00 |
| A-20k-Wav | 0,20 | 0,31 | 0,37 | 0,32 | 0,38 | 0,20 | 0,32 | 0,34 |
| A-130k-Agr | 0,20 | 0,85 | 0,90 | 0,77 | 0,96 | 0,83 | 0,82 | 0,91 |
| A-130k-LED | 0,04 | 0,08 | 0,02 | 0,09 | 0,08 | 0,02 | 0,08 | 0,03 |
| A-130k-Mix | 0,27 | 0,94 | 1,00 | 0,89 | 0,99 | 1,00 | 0,94 | 1,00 |
| A-130k-RBF | 0,13 | 0,46 | 0,73 | 0,40 | 0,67 | 0,60 | 0,45 | 0,74 |
| A-130k-Sine | 0,36 | 0,98 | 1,00 | 0,90 | 1,00 | 0,99 | 0,98 | 1,00 |
| A-130k-Wav | 0,07 | 0,25 | 0,43 | 0,20 | 0,34 | 0,35 | 0,28 | 0,38 |
| A-890k-Agr | 0,09 | 0,88 | 0,94 | 0,74 | 0,97 | 0,67 | 0,84 | 0,94 |
| A-890k-LED | 0,01 | 0,04 | 0,01 | 0,04 | 0,03 | 0,01 | 0,03 | 0,02 |
| A-890k-Mix | 0,13 | 0,94 | 1,00 | 0,89 | 0,99 | 0,98 | 0,94 | 1,00 |
| A-890k-RBF | 0,05 | 0,35 | 0,76 | 0,23 | 0,62 | 0,50 | 0,33 | 0,73 |
| A-890k-Sine | 0,22 | 0,98 | 1,00 | 0,90 | 1,00 | 1,00 | 0,98 | 1,00 |
| A-890k-Wav | 0,03 | 0,18 | 0,37 | 0,11 | 0,25 | 0,22 | 0,17 | 0,25 |
| Vitórias | 0 | 2 | 4 | 1 | 5 | 1 | 2 | 5 |
| Derrotas | 7 | 2 | 0 | 5 | 0 | 3 | 3 | 0 |
| Média Postos | 7,58 | 4,56 | 3,28 | 5,64 | 2,50 | 4,83 | 4,78 | 2,83 |

Fonte: O AUTOR (2020).

Tabela 55 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do MSDDE.

| Dataset | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|--------------|----------|-------------|-------------|-------------|-------------|-------------|----------|-------------|
| A-20k-Agr | 0,43 | 0,90 | 0,79 | 0,89 | 0,97 | 0,82 | 0,90 | 0,89 |
| A-20k-LED | 0,09 | 0,40 | 0,07 | 0,37 | 0,28 | 0,10 | 0,37 | 0,17 |
| A-20k-Mix | 0,65 | 0,99 | 1,00 | 0,99 | 1,00 | 1,00 | 0,99 | 1,00 |
| A-20k-RBF | 0,30 | 0,67 | 0,71 | 0,60 | 0,72 | 0,61 | 0,66 | 0,74 |
| A-20k-Sine | 0,62 | 0,99 | 1,00 | 0,96 | 1,00 | 0,99 | 0,99 | 1,00 |
| A-20k-Wav | 0,25 | 0,68 | 0,59 | 0,77 | 0,67 | 0,34 | 0,70 | 0,64 |
| A-130k-Agr | 0,20 | 0,66 | 0,83 | 0,59 | 0,87 | 0,60 | 0,62 | 0,90 |
| A-130k-LED | 0,04 | 0,19 | 0,11 | 0,16 | 0,26 | 0,06 | 0,16 | 0,25 |
| A-130k-Mix | 0,45 | 0,99 | 1,00 | 0,99 | 1,00 | 0,95 | 0,99 | 1,00 |
| A-130k-RBF | 0,13 | 0,36 | 0,72 | 0,32 | 0,54 | 0,32 | 0,33 | 0,64 |
| A-130k-Sine | 0,33 | 0,98 | 1,00 | 0,93 | 0,99 | 0,71 | 0,98 | 0,99 |
| A-130k-Wav | 0,08 | 0,63 | 0,65 | 0,63 | 0,68 | 0,37 | 0,67 | 0,70 |
| A-890k-Agr | 0,08 | 0,31 | 0,84 | 0,27 | 0,69 | 0,40 | 0,28 | 0,84 |
| A-890k-LED | 0,01 | 0,08 | 0,06 | 0,08 | 0,16 | 0,07 | 0,07 | 0,14 |
| A-890k-Mix | 0,18 | 0,98 | 1,00 | 0,94 | 1,00 | 0,61 | 0,98 | 1,00 |
| A-890k-RBF | 0,05 | 0,15 | 0,55 | 0,13 | 0,31 | 0,17 | 0,14 | 0,41 |
| A-890k-Sine | 0,14 | 0,90 | 0,97 | 0,79 | 0,95 | 0,46 | 0,89 | 0,96 |
| A-890k-Wav | 0,03 | 0,31 | 0,58 | 0,29 | 0,45 | 0,25 | 0,33 | 0,47 |
| Vitórias | 0 | 3 | 2 | 2 | 5 | 1 | 3 | 5 |
| Derrotas | 7 | 2 | 0 | 4 | 0 | 6 | 2 | 0 |
| Média Postos | 7,94 | 4,14 | 3,28 | 5,44 | 2,39 | 5,97 | 4,50 | 2,33 |

Fonte: O AUTOR (2020).

Tabela 56 – Médias das acurácias em percentual (%) utilizando HT. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,55 | 86,17 | 87,05 | 86,22 | 86,51 |
| A-20k-LED | 68,90 | 68,99 | 70,16 | 69,24 | 69,26 |
| A-20k-Mix | 90,82 | 90,81 | 90,82 | 90,57 | 90,81 |
| A-20k-RBF | 74,93 | 74,73 | 74,91 | 74,77 | 74,95 |
| A-20k-Sine | 89,78 | 89,76 | 89,79 | 89,65 | 89,76 |
| A-20k-Wav | 79,30 | 79,14 | 79,32 | 78,95 | 79,27 |
| A-130k-Agr | 92,03 | 91,96 | 92,01 | 91,93 | 91,96 |
| A-130k-LED | 71,19 | 71,47 | 71,72 | 71,52 | 71,54 |
| A-130k-Mix | 92,59 | 92,60 | 92,61 | 92,59 | 92,60 |
| A-130k-RBF | 79,59 | 79,73 | 79,50 | 79,71 | 79,66 |
| A-130k-Sine | 93,34 | 93,35 | 93,35 | 93,34 | 93,35 |
| A-130k-Wav | 79,79 | 79,86 | 79,83 | 79,62 | 79,78 |
| A-890k-Agr | 95,26 | 95,26 | 95,26 | 94,09 | 95,26 |
| A-890k-LED | 71,59 | 72,08 | 71,97 | 71,93 | 72,00 |
| A-890k-Mix | 94,95 | 94,97 | 94,98 | 94,97 | 94,97 |
| A-890k-RBF | 85,09 | 85,12 | 84,98 | 84,98 | 85,05 |
| A-890k-Sine | 96,96 | 96,97 | 96,97 | 96,97 | 96,97 |
| A-890k-Wav | 81,73 | 81,98 | 81,73 | 81,93 | 81,77 |
| G-20k-Agr | 84,35 | 81,55 | 85,09 | 83,35 | 83,59 |
| G-20k-LED | 68,81 | 68,93 | 69,63 | 68,98 | 69,16 |
| G-20k-Mix | 87,29 | 87,45 | 87,32 | 87,33 | 87,45 |
| G-20k-RBF | 74,00 | 73,60 | 74,08 | 73,54 | 73,83 |
| G-20k-Sine | 86,58 | 86,64 | 86,69 | 86,73 | 86,77 |
| G-20k-Wav | 78,57 | 78,35 | 78,73 | 78,31 | 78,59 |
| G-130k-Agr | 91,74 | 90,65 | 91,72 | 89,68 | 90,76 |
| G-130k-LED | 71,09 | 71,01 | 71,59 | 71,12 | 71,51 |
| G-130k-Mix | 92,04 | 92,10 | 92,07 | 92,07 | 92,10 |
| G-130k-RBF | 79,36 | 79,29 | 79,22 | 79,36 | 79,31 |
| G-130k-Sine | 92,77 | 92,82 | 92,81 | 92,66 | 92,82 |
| G-130k-Wav | 79,76 | 79,75 | 79,72 | 79,79 | 79,75 |
| G-890k-Agr | 95,23 | 93,68 | 95,23 | 92,61 | 94,17 |
| G-890k-LED | 71,59 | 72,09 | 71,95 | 71,96 | 71,99 |
| G-890k-Mix | 94,85 | 94,88 | 94,87 | 94,88 | 94,88 |
| G-890k-RBF | 85,08 | 85,04 | 84,90 | 84,79 | 85,02 |
| G-890k-Sine | 96,84 | 96,87 | 96,86 | 96,87 | 96,87 |
| G-890k-Wav | 81,69 | 81,93 | 81,69 | 81,90 | 81,69 |
| Airlines | 65,72 | 65,94 | 65,94 | 66,19 | 65,92 |
| Connect4 | 74,93 | 74,47 | 75,05 | 74,65 | 74,86 |
| Outdoor | 59,28 | 58,91 | 59,96 | 59,53 | 59,28 |
| Rialto | 35,78 | 35,49 | 38,21 | 36,16 | 36,19 |
| SpamData | 92,18 | 91,91 | 92,42 | 91,86 | 91,91 |
| WineRed | 52,27 | 51,16 | 51,16 | 51,02 | 51,16 |
| WineWhite | 44,42 | 44,46 | 44,42 | 45,13 | 44,42 |
| Vitórias | 0 | 0 | 4 | 0 | 2 |
| Derrotas | 1 | 2 | 0 | 2 | 1 |
| Média Postos | 3,33 | 3,02 | 2,44 | 3,53 | 2,67 |

Fonte: O AUTOR (2020).

Tabela 57 – Médias das acurácias em percentual (%) utilizando NB. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|--------------|--------------|--------------|--------------|--------------|
| A-20k-Agr | 86,86 | 85,64 | 86,72 | 87,06 | 86,41 |
| A-20k-LED | 67,67 | 65,15 | 67,38 | 70,68 | 65,46 |
| A-20k-Mix | 91,20 | 91,17 | 91,17 | 91,19 | 91,10 |
| A-20k-RBF | 74,90 | 74,53 | 74,83 | 74,95 | 74,64 |
| A-20k-Sine | 89,19 | 89,17 | 89,18 | 89,19 | 89,01 |
| A-20k-Wav | 79,67 | 79,57 | 79,67 | 79,69 | 79,10 |
| A-130k-Agr | 88,91 | 88,67 | 88,91 | 88,91 | 87,66 |
| A-130k-LED | 72,25 | 68,26 | 71,45 | 73,25 | 68,67 |
| A-130k-Mix | 91,94 | 91,94 | 91,94 | 91,94 | 91,86 |
| A-130k-RBF | 75,30 | 74,66 | 75,04 | 75,33 | 73,57 |
| A-130k-Sine | 89,49 | 89,49 | 89,49 | 89,49 | 89,03 |
| A-130k-Wav | 80,34 | 80,31 | 80,34 | 80,34 | 80,06 |
| A-890k-Agr | 89,36 | 89,25 | 89,36 | 89,35 | 84,50 |
| A-890k-LED | 73,65 | 71,48 | 73,43 | 73,83 | 68,45 |
| A-890k-Mix | 92,08 | 92,07 | 92,07 | 92,07 | 89,53 |
| A-890k-RBF | 75,45 | 75,17 | 75,41 | 75,45 | 71,20 |
| A-890k-Sine | 89,51 | 89,50 | 89,50 | 89,50 | 83,35 |
| A-890k-Wav | 80,40 | 80,38 | 80,40 | 80,40 | 80,02 |
| G-20k-Agr | 82,18 | 76,59 | 80,44 | 84,87 | 80,29 |
| G-20k-LED | 66,47 | 62,70 | 66,02 | 69,56 | 62,86 |
| G-20k-Mix | 87,68 | 87,81 | 87,82 | 87,70 | 87,70 |
| G-20k-RBF | 73,43 | 70,38 | 72,66 | 73,64 | 71,71 |
| G-20k-Sine | 86,60 | 86,89 | 86,89 | 86,84 | 86,85 |
| G-20k-Wav | 78,92 | 78,83 | 78,84 | 78,94 | 78,25 |
| G-130k-Agr | 88,29 | 87,47 | 87,50 | 88,43 | 86,33 |
| G-130k-LED | 71,69 | 67,35 | 70,90 | 72,61 | 65,50 |
| G-130k-Mix | 91,40 | 91,42 | 91,42 | 91,40 | 90,98 |
| G-130k-RBF | 75,02 | 74,09 | 74,77 | 75,08 | 72,91 |
| G-130k-Sine | 89,20 | 89,31 | 89,31 | 89,31 | 88,09 |
| G-130k-Wav | 80,06 | 79,95 | 80,05 | 80,07 | 79,85 |
| G-890k-Agr | 89,27 | 87,11 | 88,51 | 89,26 | 83,67 |
| G-890k-LED | 73,54 | 70,78 | 72,78 | 73,67 | 66,77 |
| G-890k-Mix | 92,00 | 92,00 | 92,00 | 92,00 | 86,97 |
| G-890k-RBF | 75,40 | 74,92 | 75,18 | 75,42 | 70,50 |
| G-890k-Sine | 89,48 | 89,51 | 89,51 | 89,50 | 80,29 |
| G-890k-Wav | 80,34 | 80,27 | 80,34 | 80,35 | 79,89 |
| Airlines | 66,95 | 66,91 | 67,05 | 67,04 | 66,57 |
| Connect4 | 74,77 | 74,07 | 74,57 | 74,95 | 74,51 |
| Outdoor | 59,84 | 59,03 | 60,71 | 60,71 | 59,64 |
| Rialto | 26,49 | 25,59 | 24,45 | 31,34 | 24,40 |
| SpamData | 91,35 | 91,32 | 91,32 | 91,58 | 91,28 |
| WineRed | 47,52 | 47,70 | 47,52 | 47,50 | 47,70 |
| WineWhite | 44,93 | 44,77 | 46,64 | 46,66 | 42,82 |
| Vitórias | 3 | 1 | 2 | 4 | 0 |
| Derrotas | 1 | 3 | 2 | 0 | 4 |
| Média Postos | 2,30 | 3,69 | 2,62 | 1,77 | 4,63 |

Fonte: O AUTOR (2020).

Tabela 58 – Médias dos MTR utilizando HT. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|-----------------|----------------|----------------|-------------|---------------|
| A-20k-Agr | 75,62 | 43,10 | 112,53 | 34,53 | 62,23 |
| A-20k-LED | 1,61 | 1,88 | 2,53 | 2,01 | 2,02 |
| A-20k-Mix | 345,36 | 310,79 | 320,09 | 156,50 | 310,79 |
| A-20k-RBF | 8,96 | 5,95 | 16,48 | 4,04 | 9,01 |
| A-20k-Sine | 290,34 | 265,14 | 287,25 | 176,32 | 265,14 |
| A-20k-Wav | 0,65 | 0,55 | 0,93 | 0,51 | 0,69 |
| A-130k-Agr | 1082,95 | 769,42 | 1011,64 | 493,86 | 791,62 |
| A-130k-LED | 1,76 | 2,40 | 10,34 | 2,53 | 2,67 |
| A-130k-Mix | 2004,67 | 1972,20 | 2100,45 | 1376,29 | 1972,20 |
| A-130k-RBF | 258,75 | 241,43 | 309,98 | 174,72 | 276,77 |
| A-130k-Sine | 2038,03 | 1722,00 | 1991,98 | 1395,62 | 1722,00 |
| A-130k-Wav | 56,59 | 4,03 | 45,90 | 9,22 | 56,81 |
| A-890k-Agr | 9273,66 | 6027,95 | 6661,40 | 2089,80 | 6027,95 |
| A-890k-LED | 2,54 | 2,04 | 7,52 | 2,37 | 4,00 |
| A-890k-Mix | 16707,50 | 13344,65 | 15368,69 | 9721,59 | 13344,65 |
| A-890k-RBF | 2164,47 | 2196,50 | 2011,13 | 927,49 | 1931,00 |
| A-890k-Sine | 16869,91 | 11848,69 | 14445,17 | 8769,93 | 11848,69 |
| A-890k-Wav | 165,06 | 71,67 | 118,63 | 100,57 | 159,78 |
| G-20k-Agr | 4,86 | 2,47 | 11,32 | 6,26 | 6,59 |
| G-20k-LED | 1,69 | 1,89 | 2,17 | 1,81 | 2,02 |
| G-20k-Mix | 5,59 | 0,93 | 0,86 | 0,97 | 0,93 |
| G-20k-RBF | 1,67 | 0,94 | 5,78 | 1,04 | 1,42 |
| G-20k-Sine | 2,72 | 4,05 | 0,98 | 7,37 | 1,38 |
| G-20k-Wav | 0,44 | 0,28 | 0,49 | 0,28 | 0,48 |
| G-130k-Agr | 110,83 | 98,88 | 160,33 | 95,00 | 109,14 |
| G-130k-LED | 1,68 | 1,88 | 3,69 | 2,33 | 2,61 |
| G-130k-Mix | 30,07 | 0,85 | 1,09 | 27,86 | 0,85 |
| G-130k-RBF | 62,05 | 20,36 | 59,36 | 43,19 | 50,23 |
| G-130k-Sine | 18,97 | 1,28 | 1,19 | 1,41 | 1,28 |
| G-130k-Wav | 3,14 | 0,67 | 3,64 | 0,46 | 3,10 |
| G-890k-Agr | 661,58 | 688,40 | 1455,29 | 467,27 | 744,16 |
| G-890k-LED | 6,83 | 2,07 | 5,22 | 1,61 | 3,68 |
| G-890k-Mix | 309,88 | 223,42 | 0,90 | 223,64 | 223,48 |
| G-890k-RBF | 556,60 | 498,08 | 467,81 | 261,18 | 504,81 |
| G-890k-Sine | 14,34 | 441,31 | 0,91 | 298,28 | 443,33 |
| G-890k-Wav | 50,25 | 6,81 | 59,91 | 33,95 | 24,31 |
| Vitórias | 3 | 1 | 3 | 0 | 2 |
| Derrotas | 0 | 3 | 0 | 4 | 2 |
| Média Postos | 2,25 | 3,82 | 2,11 | 4,04 | 2,78 |

Fonte: O AUTOR (2020).

Tabela 59 – Médias dos MTR utilizando NB. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|-----------------|----------------|---------------|---------------|----------------|
| A-20k-Agr | 97,46 | 63,24 | 86,87 | 118,69 | 64,08 |
| A-20k-LED | 0,95 | 0,48 | 0,90 | 15,05 | 0,55 |
| A-20k-Mix | 348,85 | 306,08 | 306,08 | 324,05 | 237,82 |
| A-20k-RBF | 10,13 | 3,28 | 9,93 | 12,75 | 6,64 |
| A-20k-Sine | 297,15 | 261,46 | 265,71 | 286,88 | 160,53 |
| A-20k-Wav | 31,52 | 24,98 | 31,51 | 41,05 | 6,60 |
| A-130k-Agr | 712,61 | 507,19 | 617,18 | 700,10 | 254,89 |
| A-130k-LED | 33,83 | 1,59 | 34,57 | 98,34 | 12,56 |
| A-130k-Mix | 2300,57 | 2037,77 | 2037,77 | 2099,09 | 910,46 |
| A-130k-RBF | 132,46 | 34,36 | 130,90 | 171,88 | 8,35 |
| A-130k-Sine | 2006,22 | 1742,30 | 1768,44 | 1810,44 | 509,05 |
| A-130k-Wav | 465,94 | 371,29 | 470,34 | 469,71 | 150,79 |
| A-890k-Agr | 4678,56 | 3626,46 | 4114,44 | 4124,00 | 427,01 |
| A-890k-LED | 206,45 | 64,08 | 198,39 | 358,08 | 48,87 |
| A-890k-Mix | 15114,63 | 13633,88 | 13633,88 | 13760,01 | 3255,63 |
| A-890k-RBF | 442,09 | 194,03 | 468,39 | 518,85 | 1,75 |
| A-890k-Sine | 12848,82 | 11314,39 | 11424,12 | 11429,14 | 1092,68 |
| A-890k-Wav | 1763,41 | 2086,66 | 1942,70 | 1869,03 | 633,54 |
| G-20k-Agr | 5,13 | 7,92 | 11,92 | 15,35 | 13,52 |
| G-20k-LED | 0,59 | 0,17 | 0,57 | 3,94 | 0,23 |
| G-20k-Mix | 3,58 | 1,07 | 1,07 | 0,81 | 1,07 |
| G-20k-RBF | 3,50 | 1,55 | 3,19 | 7,27 | 2,11 |
| G-20k-Sine | 2,58 | 6,93 | 4,01 | 0,94 | 3,96 |
| G-20k-Wav | 0,31 | 0,30 | 0,30 | 0,38 | 0,28 |
| G-130k-Agr | 67,66 | 126,74 | 126,42 | 151,21 | 122,45 |
| G-130k-LED | 7,60 | 0,83 | 7,80 | 31,46 | 0,59 |
| G-130k-Mix | 32,77 | 0,87 | 0,87 | 0,96 | 52,67 |
| G-130k-RBF | 21,25 | 11,84 | 18,67 | 26,34 | 12,22 |
| G-130k-Sine | 5,37 | 1,30 | 1,08 | 1,03 | 52,38 |
| G-130k-Wav | 57,31 | 40,90 | 59,05 | 64,92 | 51,70 |
| G-890k-Agr | 500,15 | 590,94 | 669,94 | 866,45 | 253,82 |
| G-890k-LED | 52,53 | 2,23 | 38,94 | 177,41 | 0,72 |
| G-890k-Mix | 223,83 | 0,79 | 0,80 | 0,81 | 1204,30 |
| G-890k-RBF | 114,11 | 49,09 | 87,79 | 136,06 | 28,66 |
| G-890k-Sine | 14,12 | 404,15 | 105,30 | 1,27 | 320,04 |
| G-890k-Wav | 252,98 | 356,81 | 254,68 | 250,53 | 345,24 |
| Vitórias | 2 | 1 | 2 | 3 | 0 |
| Derrotas | 0 | 3 | 1 | 0 | 4 |
| Média Postos | 2,33 | 3,74 | 2,90 | 1,97 | 4,06 |

Fonte: O AUTOR (2020).

Tabela 60 – Médias dos MCC utilizando HT. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|-------------|-------------|-------------|-------------|-------------|
| A-20k-Agr | 0,87 | 0,83 | 0,94 | 0,73 | 0,86 |
| A-20k-LED | 0,05 | 0,01 | 0,14 | 0,04 | 0,05 |
| A-20k-Mix | 0,96 | 1,00 | 1,00 | 1,00 | 1,00 |
| A-20k-RBF | 0,67 | 0,63 | 0,66 | 0,50 | 0,69 |
| A-20k-Sine | 0,98 | 1,00 | 0,99 | 1,00 | 1,00 |
| A-20k-Wav | 0,38 | 0,38 | 0,39 | 0,20 | 0,38 |
| A-130k-Agr | 0,96 | 0,91 | 0,97 | 0,84 | 0,92 |
| A-130k-LED | 0,01 | 0,01 | 0,05 | 0,02 | 0,03 |
| A-130k-Mix | 0,95 | 0,99 | 1,00 | 1,00 | 0,99 |
| A-130k-RBF | 0,70 | 0,73 | 0,70 | 0,55 | 0,72 |
| A-130k-Sine | 0,98 | 1,00 | 1,00 | 0,99 | 1,00 |
| A-130k-Wav | 0,38 | 0,38 | 0,36 | 0,32 | 0,39 |
| A-890k-Agr | 0,98 | 0,94 | 0,98 | 0,67 | 0,94 |
| A-890k-LED | 0,01 | 0,01 | 0,02 | 0,00 | 0,01 |
| A-890k-Mix | 0,95 | 0,99 | 0,99 | 0,99 | 0,99 |
| A-890k-RBF | 0,71 | 0,73 | 0,66 | 0,51 | 0,69 |
| A-890k-Sine | 0,98 | 1,00 | 1,00 | 1,00 | 1,00 |
| A-890k-Wav | 0,27 | 0,36 | 0,24 | 0,22 | 0,27 |
| Vitórias | 0 | 1 | 1 | 0 | 1 |
| Derrotas | 0 | 0 | 0 | 3 | 0 |
| Média Postos | 3,39 | 2,86 | 2,28 | 4,00 | 2,47 |

Fonte: O AUTOR (2020).

Tabela 61 – Médias dos MCC utilizando NB. Considerando comparações com parametrizações do MDDE.

| Dataset | MDDEC1a | MDDEC1b | MDDEC2a | MDDEC2b | MDDEC3a |
|--------------|-------------|-------------|-------------|-------------|-------------|
| A-20k-Agr | 0,89 | 0,81 | 0,90 | 0,93 | 0,83 |
| A-20k-LED | 0,15 | 0,08 | 0,17 | 0,22 | 0,09 |
| A-20k-Mix | 0,99 | 1,00 | 1,00 | 1,00 | 1,00 |
| A-20k-RBF | 0,74 | 0,71 | 0,75 | 0,75 | 0,59 |
| A-20k-Sine | 1,00 | 1,00 | 1,00 | 1,00 | 0,99 |
| A-20k-Wav | 0,66 | 0,60 | 0,66 | 0,68 | 0,37 |
| A-130k-Agr | 0,90 | 0,86 | 0,91 | 0,92 | 0,62 |
| A-130k-LED | 0,19 | 0,12 | 0,21 | 0,26 | 0,07 |
| A-130k-Mix | 0,99 | 1,00 | 1,00 | 1,00 | 0,95 |
| A-130k-RBF | 0,60 | 0,71 | 0,64 | 0,59 | 0,33 |
| A-130k-Sine | 0,99 | 1,00 | 0,99 | 0,99 | 0,70 |
| A-130k-Wav | 0,69 | 0,65 | 0,71 | 0,69 | 0,36 |
| A-890k-Agr | 0,83 | 0,85 | 0,84 | 0,80 | 0,39 |
| A-890k-LED | 0,12 | 0,06 | 0,14 | 0,16 | 0,06 |
| A-890k-Mix | 0,99 | 1,00 | 1,00 | 1,00 | 0,61 |
| A-890k-RBF | 0,38 | 0,55 | 0,41 | 0,38 | 0,19 |
| A-890k-Sine | 0,93 | 0,97 | 0,96 | 0,95 | 0,46 |
| A-890k-Wav | 0,46 | 0,59 | 0,49 | 0,47 | 0,25 |
| Vitórias | 1 | 1 | 2 | 2 | 0 |
| Derrotas | 2 | 0 | 0 | 0 | 4 |
| Média Postos | 3,28 | 2,75 | 2,06 | 2,19 | 4,72 |

Fonte: O AUTOR (2020).

D.2 RESUMO DOS TESTES DE HIPÓTESES

Tabela 62 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₂ | H ₀ | H ₀ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC2a | H ₁ | H ₀ | H ₀ | * | H ₁ | H ₀ | H ₁ | H ₀ |
| DDEC2b | H ₂ | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₂ | H ₂ |
| DDEC2c | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | * | H ₀ |
| DDEC3b | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 5 | 4 | 4 | 0 | 3 | 5 | 4 |
| $\sum H_1$ - Vitórias | 1 | 1 | 1 | 0 | 7 | 0 | 1 | 2 |
| $\sum H_2$ - Derrotas | 1 | 1 | 2 | 3 | 0 | 4 | 1 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | 0 | -1 | -3 | 7 | -4 | 0 | 1 |
| Rank | 4,0 | 4,0 | 6,0 | 7,0 | 1,0 | 8,0 | 4,0 | 2,0 |

Fonte: O AUTOR (2020).

Tabela 63 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1b | H ₁ | * | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₂ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₂ |
| DDEC3b | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | * |
| $\sum H_0$ - Empates | 3 | 0 | 0 | 3 | 3 | 0 | 3 | 0 |
| $\sum H_1$ - Vitórias | 4 | 3 | 1 | 4 | 4 | 0 | 4 | 2 |
| $\sum H_2$ - Derrotas | 0 | 4 | 6 | 0 | 0 | 7 | 0 | 5 |
| $\sum H_1 - \sum H_2$ | 4 | -1 | -5 | 4 | 4 | -7 | 4 | -3 |
| Rank | 2,5 | 5,0 | 7,0 | 2,5 | 2,5 | 8,0 | 2,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 64 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 4 | 0 | 4 | 4 | 0 | 5 | 2 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 3 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 0 | 0 | 6 | 0 | 0 | 7 | 0 | 3 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | -5 | 3 | 3 | -7 | 2 | -1 |
| Rank | 4,5 | 2,0 | 7,0 | 2,0 | 2,0 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 65 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 4 | 4 | 0 | 3 | 4 | 0 | 4 | 3 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 4 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 1 | 0 | 6 | 0 | 0 | 7 | 1 | 2 |
| $\sum H_1 - \sum H_2$ | 1 | 3 | -5 | 4 | 3 | -7 | 1 | 0 |
| Rank | 4,5 | 2,5 | 7,0 | 1,0 | 2,5 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 66 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1b | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1c | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₀ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 3 | 4 | 3 | 1 | 2 | 4 | 3 | 2 |
| $\sum H_1$ - Vitórias | 1 | 1 | 4 | 0 | 5 | 0 | 1 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 6 | 0 | 3 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -2 | -1 | 4 | -6 | 5 | -3 | -2 | 5 |
| Rank | 5,5 | 4,0 | 3,0 | 8,0 | 1,5 | 7,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 67 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do DDE.

| Dataset | DDEC1a | DDEC1b | DDEC1c | DDEC2a | DDEC2b | DDEC2c | DDEC3a | DDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC1b | H ₂ | * | H ₀ | H ₂ | H ₁ | H ₂ | H ₂ | H ₁ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₀ | * | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 2 | 1 | 6 | 1 | 2 | 0 | 2 | 2 |
| $\sum H_1$ - Vitórias | 2 | 4 | 1 | 1 | 5 | 0 | 2 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 5 | 0 | 7 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -1 | 2 | 1 | -4 | 5 | -7 | -1 | 5 |
| Rank | 5,5 | 3,0 | 4,0 | 7,0 | 1,5 | 8,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 68 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₂ | H ₀ | H ₀ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC2a | H ₁ | H ₀ | H ₀ | * | H ₁ | H ₀ | H ₁ | H ₀ |
| DDEC2b | H ₂ | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₂ | H ₂ |
| DDEC2c | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | * | H ₀ |
| DDEC3b | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 5 | 4 | 4 | 0 | 3 | 5 | 4 |
| $\sum H_1$ - Vitórias | 1 | 1 | 1 | 0 | 7 | 0 | 1 | 2 |
| $\sum H_2$ - Derrotas | 1 | 1 | 2 | 3 | 0 | 4 | 1 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | 0 | -1 | -3 | 7 | -4 | 0 | 1 |
| Rank | 4,0 | 4,0 | 6,0 | 7,0 | 1,0 | 8,0 | 4,0 | 2,0 |

Fonte: O AUTOR (2020).

Tabela 69 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1b | H ₁ | * | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₂ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₂ |
| DDEC3b | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | * |
| $\sum H_0$ - Empates | 3 | 0 | 0 | 3 | 3 | 0 | 3 | 0 |
| $\sum H_1$ - Vitórias | 4 | 3 | 1 | 4 | 4 | 0 | 4 | 2 |
| $\sum H_2$ - Derrotas | 0 | 4 | 6 | 0 | 0 | 7 | 0 | 5 |
| $\sum H_1 - \sum H_2$ | 4 | -1 | -5 | 4 | 4 | -7 | 4 | -3 |
| Rank | 2,5 | 5,0 | 7,0 | 2,5 | 2,5 | 8,0 | 2,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 70 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 4 | 0 | 4 | 4 | 0 | 5 | 2 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 3 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 0 | 0 | 6 | 0 | 0 | 7 | 0 | 3 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | -5 | 3 | 3 | -7 | 2 | -1 |
| Rank | 4,5 | 2,0 | 7,0 | 2,0 | 2,0 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 71 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 4 | 4 | 0 | 3 | 4 | 0 | 4 | 3 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 4 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 1 | 0 | 6 | 0 | 0 | 7 | 1 | 2 |
| $\sum H_1 - \sum H_2$ | 1 | 3 | -5 | 4 | 3 | -7 | 1 | 0 |
| Rank | 4,5 | 2,5 | 7,0 | 1,0 | 2,5 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 72 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1b | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1c | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₀ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 3 | 4 | 3 | 1 | 2 | 4 | 3 | 2 |
| $\sum H_1$ - Vitórias | 1 | 1 | 4 | 0 | 5 | 0 | 1 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 6 | 0 | 3 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -2 | -1 | 4 | -6 | 5 | -3 | -2 | 5 |
| Rank | 5,5 | 4,0 | 3,0 | 8,0 | 1,5 | 7,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 73 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do MSDDE.

| Detector | MSDDEC1a | MSDDEC1b | MSDDEC1c | MSDDEC2a | MSDDEC2b | MSDDEC2c | MSDDEC3a | MSDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC1b | H ₂ | * | H ₀ | H ₂ | H ₁ | H ₂ | H ₂ | H ₁ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₀ | * | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 2 | 1 | 6 | 1 | 2 | 0 | 2 | 2 |
| $\sum H_1$ - Vitórias | 2 | 4 | 1 | 1 | 5 | 0 | 2 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 5 | 0 | 7 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -1 | 2 | 1 | -4 | 5 | -7 | -1 | 5 |
| Rank | 5,5 | 3,0 | 4,0 | 7,0 | 1,5 | 8,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 74 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₂ | H ₀ | H ₀ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC2a | H ₁ | H ₀ | H ₀ | * | H ₁ | H ₀ | H ₁ | H ₀ |
| DDEC2b | H ₂ | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₂ | H ₂ |
| DDEC2c | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | * | H ₀ |
| DDEC3b | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 5 | 4 | 4 | 0 | 3 | 5 | 4 |
| $\sum H_1$ - Vitórias | 1 | 1 | 1 | 0 | 7 | 0 | 1 | 2 |
| $\sum H_2$ - Derrotas | 1 | 1 | 2 | 3 | 0 | 4 | 1 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | 0 | -1 | -3 | 7 | -4 | 0 | 1 |
| Rank | 4,0 | 4,0 | 6,0 | 7,0 | 1,0 | 8,0 | 4,0 | 2,0 |

Fonte: O AUTOR (2020).

Tabela 75 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1b | H ₁ | * | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₂ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₂ |
| DDEC3b | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₁ | * |
| $\sum H_0$ - Empates | 3 | 0 | 0 | 3 | 3 | 0 | 3 | 0 |
| $\sum H_1$ - Vitórias | 4 | 3 | 1 | 4 | 4 | 0 | 4 | 2 |
| $\sum H_2$ - Derrotas | 0 | 4 | 6 | 0 | 0 | 7 | 0 | 5 |
| $\sum H_1 - \sum H_2$ | 4 | -1 | -5 | 4 | 4 | -7 | 4 | -3 |
| Rank | 2,5 | 5,0 | 7,0 | 2,5 | 2,5 | 8,0 | 2,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 76 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₀ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₁ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 4 | 0 | 4 | 4 | 0 | 5 | 2 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 3 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 0 | 0 | 6 | 0 | 0 | 7 | 0 | 3 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | -5 | 3 | 3 | -7 | 2 | -1 |
| Rank | 4,5 | 2,0 | 7,0 | 2,0 | 2,0 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 77 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC1b | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ |
| DDEC1c | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2a | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | * | H ₀ |
| DDEC3b | H ₀ | H ₁ | H ₂ | H ₀ | H ₁ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 4 | 4 | 0 | 3 | 4 | 0 | 4 | 3 |
| $\sum H_1$ - Vitórias | 2 | 3 | 1 | 4 | 3 | 0 | 2 | 2 |
| $\sum H_2$ - Derrotas | 1 | 0 | 6 | 0 | 0 | 7 | 1 | 2 |
| $\sum H_1 - \sum H_2$ | 1 | 3 | -5 | 4 | 3 | -7 | 1 | 0 |
| Rank | 4,5 | 2,5 | 7,0 | 1,0 | 2,5 | 8,0 | 4,5 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 78 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1b | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₀ | H ₁ |
| DDEC1c | H ₂ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₂ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₀ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | * | H ₀ | H ₁ |
| DDEC3a | H ₀ | H ₀ | H ₁ | H ₂ | H ₁ | H ₀ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 3 | 4 | 3 | 1 | 2 | 4 | 3 | 2 |
| $\sum H_1$ - Vitórias | 1 | 1 | 4 | 0 | 5 | 0 | 1 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 6 | 0 | 3 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -2 | -1 | 4 | -6 | 5 | -3 | -2 | 5 |
| Rank | 5,5 | 4,0 | 3,0 | 8,0 | 1,5 | 7,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 79 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando comparações com parametrizações do MDDE.

| Detector | MDDEC1a | MDDEC1b | MDDEC1c | MDDEC2a | MDDEC2b | MDDEC2c | MDDEC3a | MDDEC3b |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DDEC1a | * | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | H ₀ | H ₁ |
| DDEC1b | H ₂ | * | H ₀ | H ₂ | H ₁ | H ₂ | H ₂ | H ₁ |
| DDEC1c | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ |
| DDEC2a | H ₁ | H ₁ | H ₀ | * | H ₁ | H ₂ | H ₁ | H ₁ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₂ | H ₂ | H ₀ |
| DDEC2c | H ₁ | * | H ₁ | H ₁ |
| DDEC3a | H ₀ | H ₁ | H ₀ | H ₂ | H ₁ | H ₂ | * | H ₁ |
| DDEC3b | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₂ | * |
| $\sum H_0$ - Empates | 2 | 1 | 6 | 1 | 2 | 0 | 2 | 2 |
| $\sum H_1$ - Vitórias | 2 | 4 | 1 | 1 | 5 | 0 | 2 | 5 |
| $\sum H_2$ - Derrotas | 3 | 2 | 0 | 5 | 0 | 7 | 3 | 0 |
| $\sum H_1 - \sum H_2$ | -1 | 2 | 1 | -4 | 5 | -7 | -1 | 5 |
| Rank | 5,5 | 3,0 | 4,0 | 7,0 | 1,5 | 8,0 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 80 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| FSDD | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ | H ₂ | H ₀ | H ₀ | H ₂ |
| ftdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₁ | H ₂ | H ₀ | H ₀ | H ₂ |
| hddma | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₁ | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ |
| DDEC1a | H ₁ | H ₀ | H ₀ | H ₁ | * | H ₁ | H ₁ | H ₂ | H ₀ | H ₀ | H ₀ |
| MSDDEC2b | H ₂ | * | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| MDDEC2a | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₂ | H ₂ | H ₂ | H ₂ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3a | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | H ₂ | * | H ₀ | H ₀ |
| ALOC3a | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | H ₂ | H ₀ | * | H ₀ |
| eDetC2a | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 4 | 6 | 6 | 3 | 5 | 1 | 2 | 0 | 5 | 5 | 3 |
| $\sum H_1$ - Vitórias | 5 | 2 | 2 | 5 | 1 | 9 | 8 | 0 | 1 | 1 | 1 |
| $\sum H_2$ - Derrotas | 1 | 2 | 2 | 2 | 4 | 0 | 0 | 10 | 4 | 4 | 6 |
| $\sum H_1 - \sum H_2$ | 4 | 0 | 0 | 3 | -3 | 9 | 8 | -10 | -3 | -3 | -5 |
| Rank | 3,0 | 5,5 | 5,5 | 4,0 | 8,0 | 1,0 | 2,0 | 11,0 | 8,0 | 8,0 | 10,0 |

Fonte: O AUTOR (2020).

Tabela 81 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ |
| fsdd | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₂ |
| ftdd | H ₀ | H ₁ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ |
| hddma | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ |
| DDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ | H ₂ |
| MSDDEC2b | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₂ |
| MDDEC2a | H ₁ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | * | H ₂ | H ₁ | H ₁ | H ₂ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3 | H ₁ | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ | H ₂ | H ₂ | * | H ₀ | H ₂ |
| ALOC3 | H ₁ | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ | H ₂ | H ₂ | H ₀ | * | H ₂ |
| eDetC2 | H ₁ | H ₂ | H ₁ | H ₁ | * |
| $\sum H_0$ - Empates | 5 | 6 | 5 | 8 | 4 | 4 | 2 | 0 | 4 | 4 | 0 |
| $\sum H_1$ - Vitórias | 5 | 4 | 2 | 2 | 6 | 6 | 2 | 0 | 3 | 3 | 1 |
| $\sum H_2$ - Derrotas | 0 | 0 | 3 | 0 | 0 | 0 | 6 | 10 | 3 | 3 | 9 |
| $\sum H_1 - \sum H_2$ | 5 | 4 | -1 | 2 | 6 | 6 | -4 | -10 | 0 | 0 | -8 |
| Rank | 3,0 | 4,0 | 8,0 | 5,0 | 1,5 | 1,5 | 9,0 | 11,0 | 6,5 | 6,5 | 10,0 |

Fonte: O AUTOR (2020).

Tabela 82 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ | H ₂ | H ₀ | H ₂ |
| fsdd | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ |
| hddma | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ |
| DDEC1a | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ |
| MSDDEC2b | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ |
| MDDEC2a | H ₀ | * | H ₂ | H ₂ | H ₀ | H ₀ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3a | H ₁ | H ₂ | * | H ₁ | H ₁ |
| ALOC3a | H ₀ | H ₂ | H ₂ | * | H ₀ |
| eDetC2a | H ₁ | H ₀ | H ₂ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 0 | 0 | 8 | 7 |
| $\sum H_1$ - Vitórias | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 2 | 2 |
| $\sum H_2$ - Derrotas | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 9 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 3 | 2 | 2 | 2 | 2 | 2 | 2 | -10 | -8 | 2 | 1 |
| Rank | 1,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 11,0 | 10,0 | 5,0 | 9,0 |

Fonte: O AUTOR (2020).

Tabela 83 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddma | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| DDEC2b | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| MSDDEC2b | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| MDDEC2a | H ₀ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | * | H ₂ | H ₀ | H ₀ | H ₀ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3 | H ₀ | H ₂ | * | H ₀ | H ₀ |
| ALOC3 | H ₀ | H ₂ | H ₀ | * | H ₀ |
| eDetC2 | H ₀ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 5 | 6 | 7 | 9 | 7 | 7 | 6 | 0 | 9 | 9 | 9 |
| $\sum H_1$ - Vitórias | 1 | 4 | 2 | 1 | 3 | 3 | 1 | 0 | 1 | 1 | 1 |
| $\sum H_2$ - Derrotas | 4 | 0 | 1 | 0 | 0 | 0 | 3 | 10 | 0 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | -3 | 4 | 1 | 1 | 3 | 3 | -2 | -10 | 1 | 1 | 1 |
| Rank | 10,0 | 1,0 | 6,0 | 6,0 | 2,5 | 2,5 | 9,0 | 11,0 | 6,0 | 6,0 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 84 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| fsdd | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| ftdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| hddma | H ₁ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₂ | H ₂ | H ₀ | H ₀ |
| DDEC1a | H ₁ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ |
| MSDDEC2b | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ |
| MDDEC2a | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₂ | H ₂ | H ₂ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3a | H ₁ | H ₂ | * | H ₁ | H ₁ |
| ALOC3a | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | * | H ₀ |
| eDetC2a | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 3 | 6 | 6 | 6 | 4 | 4 | 3 | 0 | 0 | 2 | 2 |
| $\sum H_1$ - Vitórias | 7 | 4 | 4 | 2 | 4 | 4 | 7 | 0 | 1 | 2 | 2 |
| $\sum H_2$ - Derrotas | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 10 | 9 | 6 | 6 |
| $\sum H_1 - \sum H_2$ | 7 | 4 | 4 | 0 | 2 | 2 | 7 | -10 | -8 | -4 | -4 |
| Rank | 1,5 | 3,5 | 3,5 | 7,0 | 5,5 | 5,5 | 1,5 | 11,0 | 10,0 | 8,5 | 8,5 |

Fonte: O AUTOR (2020).

Tabela 85 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de múltipla análise.

| Dataset | FHDDM | FSDD | FTDD | HDDMA | DDEC2b | MSDDEC2b | MDDEC2a | ADC3 | ALHDC3 | ALOC3 | eDetC2 |
|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ |
| fsdd | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₀ | H ₁ | H ₂ | H ₂ | H ₂ | H ₂ |
| ftdd | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₁ | H ₂ | H ₀ | H ₀ | H ₂ |
| hddma | H ₁ | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ | H ₀ |
| DDEC2b | H ₂ | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| MSDDEC2b | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ |
| MDDEC2a | H ₂ | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | * | H ₂ | H ₂ | H ₂ | H ₂ |
| ADC3 | H ₁ | * | H ₁ | H ₁ | H ₁ |
| ALHDC3 | H ₁ | H ₁ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₂ | * | H ₀ | H ₂ |
| ALOC3 | H ₁ | H ₁ | H ₀ | H ₂ | H ₁ | H ₁ | H ₁ | H ₂ | H ₀ | * | H ₂ |
| eDetC2 | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₂ | H ₁ | H ₁ | * |
| $\sum H_0$ - Empates | 2 | 3 | 6 | 1 | 3 | 4 | 2 | 0 | 2 | 2 | 1 |
| $\sum H_1$ - Vitórias | 5 | 5 | 3 | 1 | 7 | 6 | 8 | 0 | 3 | 3 | 1 |
| $\sum H_2$ - Derrotas | 3 | 2 | 1 | 8 | 0 | 0 | 0 | 10 | 5 | 5 | 8 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | 2 | -7 | 7 | 6 | 8 | -10 | -2 | -2 | -7 |
| Rank | 5,5 | 4,0 | 5,5 | 9,5 | 2,0 | 3,0 | 1,0 | 11,0 | 7,5 | 7,5 | 9,5 |

Fonte: O AUTOR (2020).

Tabela 86 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₁ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ | H ₁ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| hddma | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₁ |
| rddm | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₁ |
| wstd | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | * | H ₁ | H ₁ |
| usdd | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | * | H ₂ |
| usddx | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 10 | 10 | 10 | 7 | 6 | 8 | 6 | 3 | 3 |
| $\sum H_1$ - Vitórias | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 7 | 6 |
| $\sum H_2$ - Derrotas | 2 | 2 | 0 | 0 | 0 | 2 | 4 | 2 | 4 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | -1 | 0 | 0 | 0 | -1 | -4 | -2 | -4 | 7 | 5 |
| Rank | 4,5 | 7,5 | 4,5 | 4,5 | 4,5 | 7,5 | 10,5 | 9,0 | 10,5 | 1,0 | 2,0 |

Fonte: O AUTOR (2020).

Tabela 87 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ |
| fpdd | H ₁ | H ₀ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ |
| fsdd | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ |
| ftdd | H ₀ | H ₀ | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ |
| hddma | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ |
| hddmw | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₁ |
| rddm | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₀ |
| wstd | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | * | H ₁ | H ₁ |
| usdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ |
| usddx | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 6 | 8 | 2 | 6 | 5 | 8 | 6 | 7 | 4 | 4 | 2 |
| $\sum H_1$ - Vitórias | 2 | 0 | 0 | 3 | 1 | 1 | 2 | 3 | 0 | 6 | 8 |
| $\sum H_2$ - Derrotas | 2 | 2 | 8 | 1 | 4 | 1 | 2 | 0 | 6 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | -8 | 2 | -3 | 0 | 0 | 3 | -6 | 6 | 8 |
| Rank | 6,0 | 8,0 | 11,0 | 4,0 | 9,0 | 6,0 | 6,0 | 3,0 | 10,0 | 2,0 | 1,0 |

Fonte: O AUTOR (2020).

Tabela 88 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₁ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| hddma | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddmw | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₀ |
| rddm | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | * | H ₀ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ |
| usdd | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ |
| usddx | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 9 | 6 | 9 | 8 | 8 | 9 | 8 | 4 | 9 | 9 | 9 |
| $\sum H_1$ - Vitórias | 1 | 0 | 0 | 2 | 2 | 1 | 2 | 0 | 1 | 1 | 1 |
| $\sum H_2$ - Derrotas | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | 1 | -4 | -1 | 2 | 2 | 1 | 2 | -6 | 1 | 1 | 1 |
| Rank | 6,0 | 10,0 | 9,0 | 2,0 | 2,0 | 6,0 | 2,0 | 11,0 | 6,0 | 6,0 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 89 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₀ | H ₀ | H ₁ |
| fpdd | H ₀ | H ₀ | * | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| fsdd | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| hddma | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddmw | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ |
| rddm | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ |
| usddx | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 6 | 6 | 6 | 5 | 4 | 9 | 9 | 2 | 9 | 9 | 5 |
| $\sum H_1$ - Vitórias | 1 | 0 | 1 | 5 | 4 | 1 | 1 | 0 | 1 | 1 | 5 |
| $\sum H_2$ - Derrotas | 3 | 4 | 3 | 0 | 2 | 0 | 0 | 8 | 0 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | -2 | -4 | -2 | 5 | 2 | 1 | 1 | -8 | 1 | 1 | 5 |
| Rank | 8,5 | 10,0 | 8,5 | 1,5 | 3,0 | 5,5 | 5,5 | 11,0 | 5,5 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 90 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₁ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddma | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₀ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₁ | H ₀ |
| rddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₂ |
| usddx | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 9 | 9 | 9 | 5 | 6 | 0 | 8 | 5 | 8 |
| $\sum H_1$ - Vitórias | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 5 | 1 |
| $\sum H_2$ - Derrotas | 1 | 0 | 0 | 0 | 0 | 4 | 3 | 10 | 0 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | 1 | 1 | 1 | -3 | -2 | -10 | 2 | 5 | 0 |
| Rank | 3,5 | 2,0 | 6,0 | 6,0 | 6,0 | 10,0 | 9,0 | 11,0 | 3,5 | 1,0 | 8,0 |

Fonte: O AUTOR (2020).

Tabela 91 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₁ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddma | H ₁ | * | H ₁ | H ₂ | H ₁ | H ₁ | H ₁ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₁ | H ₀ |
| rddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₂ |
| usddx | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 8 | 8 | 8 | 0 | 5 | 0 | 8 | 5 | 7 |
| $\sum H_1$ - Vitórias | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 0 | 2 | 5 | 2 |
| $\sum H_2$ - Derrotas | 1 | 0 | 0 | 0 | 0 | 9 | 3 | 10 | 0 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | 2 | 2 | 2 | -8 | -1 | -10 | 2 | 5 | 1 |
| Rank | 5,0 | 2,0 | 5,0 | 5,0 | 5,0 | 10,0 | 9,0 | 11,0 | 5,0 | 1,0 | 8,0 |

Fonte: O AUTOR (2020).

APÊNDICE E – RESUMO DOS TESTES COM TRADICIONAIS

Tabela 92 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com HT. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₁ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ | H ₁ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| hddma | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₁ | H ₁ |
| rddm | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₁ |
| wstd | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | * | H ₁ | H ₁ |
| usdd | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | * | H ₂ |
| usddx | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₂ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 10 | 10 | 10 | 7 | 6 | 8 | 6 | 3 | 3 |
| $\sum H_1$ - Vitórias | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 7 | 6 |
| $\sum H_2$ - Derrotas | 2 | 2 | 0 | 0 | 0 | 2 | 4 | 2 | 4 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 0 | -1 | 0 | 0 | 0 | -1 | -4 | -2 | -4 | 7 | 5 |
| Rank | 4,5 | 7,5 | 4,5 | 4,5 | 4,5 | 7,5 | 10,5 | 9,0 | 10,5 | 1,0 | 2,0 |

Fonte: O AUTOR (2020).

Tabela 93 – Resumo do teste Wilcoxon signed-rank com médias de acurácias com NB. Considerando a métrica Acurácia e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _c | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₁ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ | H ₁ |
| fpdd | H ₁ | H ₀ | * | H ₁ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | H ₁ | H ₁ |
| fsdd | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ |
| ftdd | H ₀ | H ₀ | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ |
| hddma | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₁ |
| hddmw | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₁ |
| rddm | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₀ |
| wstd | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₀ | H ₁ | H ₁ | * | H ₁ | H ₁ |
| usdd | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ |
| usddx | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | * |
| $\sum H_0$ - Empates | 6 | 8 | 2 | 6 | 5 | 8 | 6 | 7 | 4 | 4 | 2 |
| $\sum H_1$ - Vitórias | 2 | 0 | 0 | 3 | 1 | 1 | 2 | 3 | 0 | 6 | 8 |
| $\sum H_2$ - Derrotas | 2 | 2 | 8 | 1 | 4 | 1 | 2 | 0 | 6 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | 0 | -2 | -8 | 2 | -3 | 0 | 0 | 3 | -6 | 6 | 8 |
| Rank | 6,0 | 8,0 | 11,0 | 4,0 | 9,0 | 6,0 | 6,0 | 3,0 | 10,0 | 2,0 | 1,0 |

Fonte: O AUTOR (2020).

Tabela 94 – Resumo do teste Wilcoxon signed-rank com médias de MTR com HT. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₁ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₂ | H ₂ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ |
| hddma | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddmw | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₀ | H ₀ |
| rddm | H ₁ | H ₀ | H ₀ | H ₁ | H ₀ | H ₁ | H ₁ | * | H ₀ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ |
| usdd | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ |
| usddx | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 9 | 6 | 9 | 8 | 8 | 9 | 8 | 4 | 9 | 9 | 9 |
| $\sum H_1$ - Vitórias | 1 | 0 | 0 | 2 | 2 | 1 | 2 | 0 | 1 | 1 | 1 |
| $\sum H_2$ - Derrotas | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | 1 | -4 | -1 | 2 | 2 | 1 | 2 | -6 | 1 | 1 | 1 |
| Rank | 6,0 | 10,0 | 9,0 | 2,0 | 2,0 | 6,0 | 2,0 | 11,0 | 6,0 | 6,0 | 6,0 |

Fonte: O AUTOR (2020).

Tabela 95 – Resumo do teste Wilcoxon signed-rank com médias de MTR com NB. Considerando a métrica MTR e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| fhddms | H ₀ | * | H ₀ | H ₁ | H ₁ | H ₀ | H ₁ | H ₀ | H ₀ | H ₀ | H ₁ |
| fpdd | H ₀ | H ₀ | * | H ₁ | H ₁ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| fsdd | H ₂ | H ₂ | H ₂ | * | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₂ | H ₂ | H ₂ | H ₁ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₁ |
| hddma | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddmw | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ |
| rddm | H ₁ | H ₀ | H ₁ | H ₁ | H ₁ | H ₁ | H ₀ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₀ | H ₀ | H ₂ | H ₀ | * | H ₀ |
| usddx | H ₂ | H ₂ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | * |
| $\sum H_0$ - Empates | 6 | 6 | 6 | 5 | 4 | 9 | 9 | 2 | 9 | 9 | 5 |
| $\sum H_1$ - Vitórias | 1 | 0 | 1 | 5 | 4 | 1 | 1 | 0 | 1 | 1 | 5 |
| $\sum H_2$ - Derrotas | 3 | 4 | 3 | 0 | 2 | 0 | 0 | 8 | 0 | 0 | 0 |
| $\sum H_1 - \sum H_2$ | -2 | -4 | -2 | 5 | 2 | 1 | 1 | -8 | 1 | 1 | 5 |
| Rank | 8,5 | 10,0 | 8,5 | 1,5 | 3,0 | 5,5 | 5,5 | 11,0 | 5,5 | 5,5 | 1,5 |

Fonte: O AUTOR (2020).

Tabela 96 – Resumo do teste Wilcoxon signed-rank com médias de MCC com HT. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₁ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddma | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₁ | H ₁ | H ₀ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₁ | H ₀ |
| rddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₂ |
| usddx | H ₀ | H ₀ | H ₂ | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 9 | 9 | 9 | 5 | 6 | 0 | 8 | 5 | 8 |
| $\sum H_1$ - Vitórias | 3 | 3 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 5 | 1 |
| $\sum H_2$ - Derrotas | 1 | 0 | 0 | 0 | 0 | 4 | 3 | 10 | 0 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | 1 | 1 | 1 | -3 | -2 | -10 | 2 | 5 | 0 |
| Rank | 3,5 | 2,0 | 6,0 | 6,0 | 6,0 | 10,0 | 9,0 | 11,0 | 3,5 | 1,0 | 8,0 |

Fonte: O AUTOR (2020).

Tabela 97 – Resumo do teste Wilcoxon signed-rank com médias de MCC com NB. Considerando a métrica MCC e comparações com os métodos propostos de única análise.

| Dataset | FHDDM | FHDDMS | FPDD | FSDD | FTDD | HDDM _a | HDDM _w | RDDM | WSTD | USDD | USDDX |
|-----------------------|----------------|----------------|----------------|----------------|----------------|-------------------|-------------------|----------------|----------------|----------------|----------------|
| FHDDM | * | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₁ | H ₀ |
| fhddms | H ₀ | * | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | H ₀ | H ₀ |
| fpdd | H ₀ | H ₀ | * | H ₀ | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| fsdd | H ₀ | H ₀ | H ₀ | * | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| ftdd | H ₀ | H ₀ | H ₀ | H ₀ | * | H ₂ | H ₀ | H ₂ | H ₀ | H ₀ | H ₀ |
| hddma | H ₁ | * | H ₁ | H ₂ | H ₁ | H ₁ | H ₁ |
| hddmw | H ₁ | H ₁ | H ₀ | H ₀ | H ₀ | H ₂ | * | H ₂ | H ₀ | H ₁ | H ₀ |
| rddm | H ₁ | H ₁ | * | H ₁ | H ₁ | H ₁ |
| wstd | H ₀ | H ₂ | H ₀ | H ₂ | * | H ₀ | H ₀ |
| usdd | H ₂ | H ₀ | H ₀ | H ₀ | H ₀ | H ₂ | H ₂ | H ₂ | H ₀ | * | H ₂ |
| usddx | H ₀ | H ₂ | H ₀ | H ₂ | H ₀ | H ₁ | * |
| $\sum H_0$ - Empates | 6 | 7 | 8 | 8 | 8 | 0 | 5 | 0 | 8 | 5 | 7 |
| $\sum H_1$ - Vitórias | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 0 | 2 | 5 | 2 |
| $\sum H_2$ - Derrotas | 1 | 0 | 0 | 0 | 0 | 9 | 3 | 10 | 0 | 0 | 1 |
| $\sum H_1 - \sum H_2$ | 2 | 3 | 2 | 2 | 2 | -8 | -1 | -10 | 2 | 5 | 1 |
| Rank | 5,0 | 2,0 | 5,0 | 5,0 | 5,0 | 10,0 | 9,0 | 11,0 | 5,0 | 1,0 | 8,0 |

Fonte: O AUTOR (2020).