



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

HUGO DE MORAES RAMOS BORBA

**DESENVOLVIMENTO E AVALIAÇÃO DO ALGORITMO PREPPY: UMA NOVA
ABORDAGEM PARA INDUÇÃO DE REGRAS DO DRSA**

RECIFE

2025

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

HUGO DE MORAES RAMOS BORBA

**DESENVOLVIMENTO E AVALIAÇÃO DO ALGORITMO PREPPY: UMA NOVA
ABORDAGEM PARA INDUÇÃO DE REGRAS DO DRSA**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do grau em Engenharia de Produção.

Orientador(a): Caroline Maria de Miranda Mota

RECIFE

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Borba, Hugo de Moraes Ramos.

Desenvolvimento e avaliação do algoritmo PREPPY: Uma nova abordagem para indução de regras do DRSA / Hugo de Moraes Ramos Borba. - Recife, 2025. 81 p. : il., tab.

Orientador(a): Caroline Maria de Miranda Mota

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Engenharia de Produção -Bacharelado, 2025.

Inclui referências, apêndices.

1. Análise de Decisão Multicritério (MCDA). 2. Dominance-based Rough Set Approach (DRSA). 3. Sistema da Informação (SI). 4. Algoritmos. 5. Rough Set Theory. 6. Indução de Regras de Classificação. I. Mota, Caroline Maria de Miranda. (Orientação). II. Título.

620 CDD (22.ed.)

HUGO DE MORAES RAMOS BORBA

**DESENVOLVIMENTO E AVALIAÇÃO DO ALGORITMO PREPPY: UMA NOVA
ABORDAGEM PARA INDUÇÃO DE REGRAS DO DRSA**

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia de
Produção da Universidade Federal de
Pernambuco, como requisito parcial para a
obtenção do grau em Engenharia de Produção.

Aprovado em: 03/04/2025.

BANCA EXAMINADORA

Prof.^a Dr.^a Caroline Maria de Miranda Mota (Orientadora)
Universidade Federal de Pernambuco

Prof.^a. Dr.^a Suzana de França Dantas Daher (Examinadora Interno)
Universidade Federal de Pernambuco

Prof. Dr. Anderson Lucas Carneiro de Lima da Silva (Examinador Interno)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradeço primeiramente aos meus pais, Miguel e Fabíola, que sempre me proporcionaram tudo o que precisei para dedicar-me intensamente à educação e ao meu desenvolvimento pessoal. Sem o apoio, os recursos e o incentivo constante de vocês, eu não teria chegado até aqui. À minha namorada, Letícia, meu agradecimento especial por estar sempre ao meu lado, acreditando em mim mais do que qualquer outra pessoa e por me inspirar a seguir em frente, mesmo nos momentos mais desafiadores. Também expresso minha sincera gratidão à minha orientadora, Caroline Mota, que dispendeu seu tempo e conhecimento para me orientar e apoiar no meu desenvolvimento acadêmico, e aos meus amigos, que sempre estiveram presentes para me motivar e compartilhar os momentos de alegria e superação. Cada um de vocês foi fundamental nessa jornada, e sou eternamente grato por todo o carinho e suporte que recebi ao longo dessa trajetória. Gostaria de fazer um agradecimento especial à minha avó Erleide, que faleceu durante o processo de escrita deste TCC, e às minhas avós e avôs, Edleusa, Gislaine e João, que continuam abençoando a minha vida com sua presença e amor.

RESUMO

Neste trabalho, apresenta-se o algoritmo PREPPY (*Permutation-based Recall and Precision Pattern Yielding*) para indução de regras de classificação, concebido como uma alternativa inovadora aos métodos tradicionais baseados em RST. Alinhado aos avanços teóricos da *Dominance-based Rough Set Approach*, o PREPPY integra critérios e atributos categóricos – ordinais ou nominais –, utilizando a estratégia de exploração de combinações de atributos e indução incremental de regras. A otimização equilibrada entre *recall* e *precision* é visada na geração de modelos concisos, interpretáveis e generalizáveis para ambientes com dados heterogêneos. A eficácia do método foi avaliada no conjunto de dados real 1985 *Auto Imports* e em comparações com o DomLEM utilizando conjuntos sintéticos, considerando métricas de classificação diversas. Os resultados evidenciam uma performance positiva no conjunto de dados real e desempenho superior em relação ao algoritmo DomLEM nos conjuntos considerados, embora apontem desafios na discriminação de classes intermediárias e na complexidade computacional.

ABSTRACT

This work presents the PREPPY algorithm (Permutation-based Recall and Precision Pattern Yielding) for inducing classification rules, conceived as an innovative alternative to traditional RST-based methods. Aligned with the theoretical advances of the Dominance-based Rough Set Approach, PREPPY integrates criteria and categorical attributes – ordinal or nominal – by employing a strategy that explores attribute combinations and implements incremental rule induction. The aim is to achieve a balanced optimization between recall and precision in the generation of concise, interpretable, and generalizable models for environments with heterogeneous data. The method's efficacy has been evaluated using the real dataset 1985 Auto Imports and through comparisons with the DomLEM algorithm on synthetic datasets, considering various classification metrics. The results reveal positive performance on the real dataset and superior performance compared to the DomLEM algorithm on the datasets considered, although challenges remain in discriminating intermediate classes and managing computational complexity.

SUMÁRIO

1. INTRODUÇÃO	9
1.1. Apresentação do Tema – Contexto e Descrição do Problema	10
1.2. Justificativa e Relevância	11
1.3. Objetivos	12
1.4. Contribuições e Impactos	13
1.5. Descrição Geral do Fluxo Metodológico	13
1.6. Estrutura do Trabalho	14
2. FUNDAMENTAÇÃO TEÓRICA E REVISÃO DA LITERATURA	16
2.1. <i>Rough Set Theory</i>	16
2.2. IRSA	20
2.3. DRSA	21
2.4. DomLEM	25
2.5. VC-IRSA E VC-DRSA	26
2.6. Avanços em <i>Rough Sets</i> e Suas Aplicações	34
3. PREPPY	38
3.1. Descrição Conceitual das Etapas	40
3.2. Métricas de Classificação	45
3.3. Implementação do PREPPY	47
3.4. Conjunto de Exemplo	49
4. AVALIAÇÃO DO PREPPY	56
4.1. Aplicação do PREPPY	57
4.1.1. O Conjunto de Dados 1985 <i>Auto Imports</i>	57
4.1.2. Apresentação dos Resultados Experimentais	60
4.1.3. Análise e Discussão dos Resultados	65
4.2. Avaliação Pela Comparação Entre PREPPY e DomLEM	67
4.2.1. Os Conjuntos de Dados Sintéticos	67
4.2.2. Procedimentos de Comparação	68

4.2.3. Apresentação e Discussão das Métricas	69
5. CONCLUSÃO	72
5.1. Síntese dos Principais Resultados e Contribuições	72
5.2. Limitações do Estudo	73
5.3. Sugestões para Trabalhos Futuros e Considerações Finais	74
REFERÊNCIAS	76
APÊNDICE	80

1. INTRODUÇÃO

A crescente complexidade dos processos decisórios em ambientes com dados heterogêneos impulsiona o desenvolvimento de métodos que integrem múltiplos critérios para extrair conhecimento (BLASZCZYNSKI et al., 2022). Nesse cenário, algoritmos de indução de regras como o PREPPY são essenciais, pois lidam com atributos diversos e induzem regras de forma incremental, equilibrando os indicadores de *recall* e *precision*. *Recall* mede a capacidade do modelo de identificar corretamente todas as instâncias positivas, calculado por:

$$[1] \quad \textit{Recall} = \frac{(\textit{Verdadeiros Positivos})}{(\textit{Verdadeiros Positivos} + \textit{Falsos Negativos})}$$

Precision avalia a qualidade das predições positivas, ou seja, a proporção de casos classificados como positivos que são realmente corretos, definida por:

$$[2] \quad \textit{Precision} = \frac{\textit{Verdadeiros Positivos}}{\textit{Verdadeiros Positivos} + \textit{Falsos Positivos}}$$

Esse equilíbrio entre *recall* e *precision* permite que o PREPPY gere regras concisas, generalizáveis e interpretáveis, contribuindo para análises decisórias mais eficazes. Sua avaliação foi realizada através da aplicação à análise de risco em automóveis, tomando como base o conjunto de dados 1985 *Auto Imports* (SCHLIMMER, 1985). Além disso, sua avaliação também inclui uma comparação empírica com o DomLEM – o algoritmo de indução de regras de classificação legado baseado em dominância – através de métricas de classificação.

Com o avanço do processamento computacional e o aumento exponencial da quantidade e da diversidade dos dados disponíveis, torna-se imperativo o desenvolvimento de modelos que possibilitem a extração de regras decisórias, capazes de auxiliar de forma precisa na tomada de decisão (DO COUTO e RANGEL, 2024). Métodos fundamentados na *Rough Set Theory* têm se mostrado especialmente úteis para essa finalidade, uma vez que permitem a identificação de padrões e relações de dominância que se ajustam tanto a contextos ordinais quanto a contextos com atributos qualitativos (BLASZCZYNSKI et al., 2022). Historicamente, os métodos como LEM e LEM2, baseados na *Indiscernibility-based Rough Set Approach* (IRSA) marcaram os primeiros passos na obtenção de subconjuntos mínimos de atributos discriminantes, evoluindo posteriormente para abordagens mais robustas como a *Dominance-based Rough Set Approach* (DRSA) com o DomLEM (GRZYMALA-BUSSE, 1992, 1997; GRECO et al., 2000). Em paralelo, a ampliação das aplicações desses métodos para domínios diversos – desde sustentabilidade empresarial (DO COUTO e RANGEL, 2024) até

prognósticos médicos (GIL-HERRERA, 2015) e diagnósticos industriais (LEZANSKI e PILACINSKA, 2016) – demonstra o potencial e a versatilidade das técnicas baseadas em *Rough Sets* em capturar as nuances e as complexidades inerentes aos dados reais.

O algoritmo PREPPY integra e aprimora essas abordagens, adaptando mecanismos de ordenação e permutação de critérios para induzir regras de forma incremental. Tais regras visam não somente representar com precisão as relações subjacentes aos dados, mas também proporcionar um equilíbrio entre *recall* e *precision*, o que é fundamental para a confiabilidade dos resultados em contextos aplicados, como na avaliação de risco de automóveis.

1.1. Apresentação do Tema – Contexto e Descrição do Problema

A indução de regras de classificação é uma ferramenta fundamental para compreender e representar o conhecimento extraído de bases de dados tabulados. Técnicas pioneiras, como o *Learning from Examples Module* (LEM), foram criadas para identificar subconjuntos mínimos de atributos capazes de discriminar classes distintas (GRZYMALA-BUSSE, 1992, 1997). Esse conceito evoluiu para o LEM2, que aprimorou a simplicidade e a interpretabilidade das regras através de coberturas locais, consolidando os fundamentos para os métodos de indução que se desenvolveriam posteriormente.

Com o aumento da necessidade de representar relações de ordem e preferência em ambientes com múltiplos critérios, as abordagens fundamentadas na *Rough Set Theory* ganharam destaque. A *Dominance-based Rough Set Approach* (DRSA) emergiu ao incorporar relações de dominância entre as instâncias, facilitando a extração de preferências a partir dos dados (PAWLAK, 1982; GRZYMALA-BUSSE, 1992; GRECO et al., 2000). Evoluindo nesse contexto, o DomLEM foi desenvolvido para induzir regras com base nessas relações, tornando-se especialmente útil quando atributos ordenados são cruciais para a estrutura dos dados (GRECO et al., 2000; KATO, ITSUNO e SAEKI, 2017).

Embora métodos como o DomLEM tenham sido aplicados com sucesso em diversas áreas, desafios persistem, principalmente relacionados à elevada complexidade computacional na análise dos *reducts* (KOMOROWSKI, POLKOWSKI e SKOWRON, 1999; PAWLAK e KOWRON, 2007). Além disso, a dificuldade em lidar com dados redundantes, inconsistentes e incertos evidencia a necessidade de soluções que sejam compatíveis com as

exigências de processamento e interpretação, especialmente quando os dados envolvem atributos tanto numéricos quanto categóricos.

Nesse cenário, o algoritmo PREPPY foi desenvolvido como uma alternativa inovadora para superar as limitações dos métodos anteriores. Ao combinar a capacidade de tratar critérios com a habilidade de integrar atributos categóricos, o PREPPY adota uma abordagem incremental que gera regras concisas e equilibradas entre *recall* e *precision*. Essa estratégia torna o modelo especialmente robusto em aplicações práticas, como na avaliação de risco de automóveis utilizando o conjunto de dados 1985 *Auto Imports* (SCHLIMMER, 1985), demonstrando seu desempenho em um contexto real. Ademais, o PREPPY foi concebido com uma estrutura genérica e flexível que o torna apto a ser aplicado em diversos outros cenários decisórios – desde avaliação de desempenho e classificação de investimentos até análise de riscos em diferentes setores – sem que sejam necessárias alterações no algoritmo, o que reforça sua adaptabilidade e eficácia na melhoria da representatividade e interpretabilidade das regras decisórias em contextos complexos.

1.2. Justificativa e Relevância

A modernização dos algoritmos de indução de regras é imperativa para acompanhar o dinamismo e a complexidade dos ambientes decisórios atuais. O desenvolvimento do PREPPY justifica-se pela necessidade premente de integrar, de forma harmoniosa e eficiente, atributos numéricos e categóricos aos processos decisórios, permitindo, assim, uma análise mais rica e detalhada dos dados. A capacidade de tratar critérios de avaliação com diferentes naturezas – desde variáveis de ganho ou custo até atributos qualitativos – representa um avanço teórico e prático que se alinha com as recomendações de diversos estudos na área (GRZYMALA-BUSSE, 1992; GRECO, MATARAZZO e SLOWINSKI, 2001).

A relevância deste estudo é evidenciada pela aplicabilidade do algoritmo em um cenário real e desafiador: a classificação de risco de automóveis. A complexidade intrínseca deste problema, que envolve múltiplas variáveis inter-relacionadas e a necessidade de balancear precisão e generalização dos resultados, torna o PREPPY uma alternativa promissora para superar limitações observadas em métodos tradicionais. Ao implementar técnicas que possibilitam a indução incremental de regras, o algoritmo não somente aumenta a acurácia das previsões como também contribui para a criação de modelos interpretáveis, em que a origem das decisões pode ser rastreada e validada pelos usuários.

Sua significância encontra suporte também na literatura especializada. Estudos que investigam os fundamentos da Rough Set Theory e suas aplicações em análises decisórias demonstram a efetividade de métodos como o DRSA e o DomLEM na extração de conhecimento a partir de dados complexos (PAWLAK, 1982; GRECO et al., 2000). Entretanto, a proposta do PREPPY introduz um diferencial importante ao incorporar nativamente os atributos categóricos e explorar de maneira inovadora a ordenação dos atributos, o que possibilita uma indução mais refinada e equilibrada das regras. Essa inovação abre novas perspectivas para o desenvolvimento de algoritmos de classificação que sejam mais adaptáveis aos desafios impostos por dados heterogêneos, contribuindo significativamente para o avanço da área.

1.3. Objetivos

O objetivo geral é desenvolver o algoritmo PREPPY compatível com uma gama diversa de atributos para a indução incremental de regras de classificação baseadas em DRSA. Já os objetivos específicos são:

- Apresentar e formalizar o *PREPPY*, descrevendo detalhadamente seus fundamentos, etapas de processamento e as métricas de classificação utilizadas (como *recall*, *precision*, *accuracy*, *f1-score*, *Matthews Correlation Coefficient* – MCC – e *Cohen's Kappa*). Essa etapa visa estabelecer uma base teórica sólida, alinhada às metodologias de *Rough Sets* e DRSA, discutidas na fundamentação teórica;
- Implementar e avaliar o PREPPY para a extração de regras decisórias capazes de classificar, de forma robusta, o risco em automóveis, utilizando o conjunto de dados 1985 *Auto Imports* (SCHLIMMER, 1985). Esta aplicação tem o intuito de demonstrar a eficácia do algoritmo no tratamento de dados heterogêneos e na indução incremental de regras que atendam aos requisitos decisórios práticos;
- Realizar uma análise comparativa entre o desempenho do PREPPY e o algoritmo legado DomLEM, através da aplicação de conjuntos de dados simulados que abrangem cenários diversos. A avaliação será baseada nas métricas de classificação estabelecidas, permitindo identificar os pontos fortes e as limitações de cada abordagem, além de evidenciar as inovações trazidas pelo PREPPY em termos de eficiência, interpretabilidade e generalização dos resultados.

1.4. Contribuições e Impactos

A proposta apresentada neste trabalho possui contribuições que se estendem tanto do ponto de vista teórico quanto prático. Em termos teóricos, a formalização do algoritmo PREPPY amplia o escopo das técnicas de *Rough Set Theory*, ao integrar de maneira eficaz atributos numéricos e categóricos na indução de regras. Essa integração é particularmente relevante para contextos em que os dados apresentam características heterogêneas, exigindo abordagens que não sacrifiquem a riqueza informacional em prol da simplicidade modelar (GRECO, MATARAZZO e SLOWINSKI, 2001).

Na prática, o impacto do PREPPY se reflete na sua aplicabilidade a problemas reais de classificação, como a avaliação de risco de automóveis. Ao permitir a indução incremental de regras que equilibram recall e precision, o algoritmo oferece resultados que são ao mesmo tempo concisos e generalizáveis. Essa característica facilita a interpretação dos modelos pelos decisores, contribuindo para a adoção de práticas mais precisas e confiáveis nos processos de definição de prêmios de seguro e na mitigação de riscos associados.

Adicionalmente, o estudo realiza uma comparação detalhada entre o desempenho do PREPPY e o método tradicional DomLEM, evidenciando a capacidade do novo algoritmo de superar limitações anteriores, mesmo que, eventualmente, venha a requerer maior poder computacional. A análise dos resultados experimentais, apoiada em métricas robustas e no confronto com cenários sintéticos e reais, reforça a eficácia do PREPPY, incentivando sua aplicação em áreas diversas que demandam métodos precisos para a extração de conhecimento e apoio à decisão.

O alcance esperado deste trabalho, portanto, transcende a mera comparação empírica entre algoritmos, contribuindo para o avanço dos estudos sobre DRSA e promovendo inovações que podem influenciar tanto o meio acadêmico quanto o setor empresarial, sobretudo no que tange à análise de dados complexos e heterogêneos.

1.5. Descrição Geral do Fluxo Metodológico

O estudo seguiu um fluxo metodológico integrado, iniciando com uma revisão teórica aprofundada sobre *Rough Set Theory*, métodos de indução de regras baseados em indiscernibilidade – IRSA com LEM e LEM2 – e abordagens de dominância – DRSA com DomLEM –, para identificar limitações dos métodos tradicionais e embasar o desenvolvimento de uma abordagem aprimorada.

Com base nessa revisão, foi desenvolvido o algoritmo PREPPY, que integra atributos numéricos e categóricos por meio da ordenação incremental dos critérios, buscando otimizar o equilíbrio entre *recall* e *precision*. A implementação do PREPPY seguiu boas práticas de engenharia de software, incluindo documentação e testes para garantir a replicabilidade dos resultados.

Em seguida, realizou-se uma preparação rigorosa dos dados do conjunto *1985 Auto Imports* (SCHLIMMER, 1985) – proveniente de diversas fontes, como o *Ward's Automotive Yearbook* e relatórios de seguradoras – envolvendo limpeza, normalização, correção de inconsistências, tratamento de valores ausentes e conversão de variáveis –, a fim de refletir fielmente os fatores que influenciam a classificação de risco.

O PREPPY foi então aplicado para extrair regras decisórias e classificar o risco dos automóveis, enquanto experimentos comparativos com o DomLEM foram conduzidos em cenários com dados sintéticos. A avaliação baseou-se em métricas consolidadas, permitindo uma análise quantitativa das abordagens, além de discutir sua interpretabilidade, robustez e tempo de execução.

Destaca-se ainda que o DomLEM foi utilizado na forma implementada no jMAF, o que exigiu a elaboração de scripts específicos: um para converter o arquivo de entrada ao formato compatível com o jMAF; outro para transformar a saída para uso no deployment; e um terceiro para o deployment das regras, que também gera a matriz de confusão e as métricas subsequentes.

1.6. Estrutura do Trabalho

Este TCC está organizado em cinco capítulos que conduzem o leitor desde a contextualização teórica até a aplicação e avaliação prática do algoritmo proposto, proporcionando uma visão abrangente da pesquisa:

- Capítulo 1 – Introdução: Apresenta o tema, o contexto, a justificativa e os objetivos do estudo, além de descrever a relevância das contribuições e o fluxo metodológico adotado. Esse capítulo estabelece a base conceitual e prática que orienta todo o trabalho.
- Capítulo 2 – Fundamentação Teórica e Revisão da Literatura: Explora os conceitos e evoluções da *Rough Set Theory* e suas aplicações em métodos de indução de regras. São discutidas as abordagens clássicas e as que utilizam relações de dominância, bem como

as extensões que introduzem consistência variável, evidenciando os avanços e as implicações práticas desses métodos para a análise decisória.

- Capítulo 3 – PREPPY: Detalha a proposta do algoritmo PREPPY, explicando a lógica e as etapas de sua implementação, que envolvem a organização dos dados, a indução de regras e a utilização de técnicas para otimizar a classificação por meio de métricas específicas. O capítulo também apresenta um exemplo prático que ilustra a aplicação do método.
- Capítulo 4 – Avaliação do PREPPY: Descreve os testes e a validação empírica do algoritmo, inicialmente aplicado a um conjunto de dados real relacionado à classificação de risco veicular e, em seguida, comparado com o algoritmo legado DomLEM por meio de experimentos com dados sintéticos. São apresentados e discutidos os resultados obtidos, destacando o desempenho do algoritmo em cenários reais e simulados.
- Capítulo 5 – Conclusão: Resume os principais resultados e contribuições do estudo, discute as limitações encontradas e propõe sugestões para trabalhos futuros, reforçando o potencial do PREPPY na aplicação de métodos de decisão multicritério e na otimização de processos de classificação.

Além desses capítulos, o trabalho conta com as referências bibliográficas e um apêndice com material complementar que auxilia na compreensão dos processos e resultados apresentados.

2. FUNDAMENTAÇÃO TEÓRICA E REVISÃO DA LITERATURA

Nas últimas décadas, os *Rough Set Theory* percorreram uma trajetória de inovação, partindo dos fundamentos de Pawlak (1982) até as abordagens de consistência variável (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011). Inicialmente desenvolvida para tratar a imprecisão dos dados sem pressuposições, a teoria evoluiu com métodos como IRSA e DRSA, que incorporam os conceitos de indiscernibilidade (GRZYMALA-BUSSE, 1992, 1997) e dominância (GRECO, MATARAZZO e SLOWINSKI, 2001) para induzir regras decisórias robustas. Algoritmos como LERS (GRZYMALA-BUSSE, 1992, 1997) e DomLEM (GRECO et al., 2000) exemplificam esse compromisso com a *precision* e a aplicação prática na modelagem de dados complexos.

Este capítulo procura aprofundar os fundamentos teóricos e as evoluções metodológicas dos Rough Sets, abrangendo desde as contribuições pioneiras de Pawlak (1982) até os avanços recentes em consistência variável, como VC-IRSA e VC-DRSA. Serão apresentados os principais conceitos, algoritmos e técnicas de indução de regras, destacando as capacidades de modelagem e a aplicabilidade prática desses métodos, bem como os desafios inerentes ao tratamento de dados imprecisos. A análise detalhada permitirá identificar as potencialidades desses modelos para a tomada de decisão e extração de conhecimento, estabelecendo uma base teórica robusta para as discussões e aplicações subsequentes.

2.1. *Rough Set Theory*

No início dos anos 1980, o pesquisador polonês Zdzisław Pawlak, revolucionou o tratamento da imprecisão dos dados ao propor uma metodologia que, ao contrário dos métodos probabilísticos, baseia-se unicamente na informação disponível, sem a necessidade de conhecimento prévio ou suposições adicionais. Essa abordagem é conhecida como *Rough Sets* (PAWLAK, 1982). Nesta teoria, considera-se um universo U de objetos em análise descrito por atributos A . Onde para qualquer subconjunto P de A , define-se uma relação de indiscernibilidade:

$$[3] \quad I_P = \{(x, y) \in U \times U \mid \forall_a \in P, a(x) = a(y)\}$$

a qual particiona o universo U em **classes de equivalência** – também denominadas conjuntos elementares. Dessa forma, dois objetos x e y pertencem à mesma classe se forem indistinguíveis com base nos atributos contidos em P (PAWLAK, 1982, 1991).

Essas classes de equivalência são também chamadas de grânulos de conhecimento. Em outras palavras, um grânulo de conhecimento representa o conjunto de todos os objetos que compartilham exatamente a mesma informação descritiva em relação ao conjunto de atributos P . Estes grânulos constituem a unidade fundamental sobre a qual são construídas as aproximações em RST, permitindo a captura do conhecimento disponível sem recorrer a informações adicionais ou a pressupostos probabilísticos (PAWLAK, 1982, 1991).

Em *Rough Sets*, a categoria de objetos que compartilham certas características – isto é, o conceito que desejamos representar – é representada por X . Entretanto, devido à imprecisão dos dados, esse conceito não pode ser definido de forma exata. Assim, recorreremos a aproximações: a **aproximação inferior** reúne os objetos que certamente pertencem a X , enquanto a **aproximação superior** inclui aqueles que possivelmente pertencem a X .

Com base na estrutura (U, IP) , definimos aproximação inferior e superior de um conjunto X de U como, respectivamente:

$$[4] \quad \underline{X} = \{x \in U \mid [x]_P \subseteq X\}$$

$$[5] \quad \overline{X} = \{x \in U \mid [x]_P \cap X \neq \emptyset\}$$

em que $[x]_P$ representa a classe de equivalência de x sob IP . Na aproximação inferior, todos os objetos da classe de equivalência de x devem estar em X . Já na aproximação superior, todos os objetos de sua classe de equivalência tem uma interseção não vazia com X , ou seja, cuja classe de equivalência possui pelo menos um elemento em X (PAWLAK, 1982, 1991).

Também segundo Pawlak (1982, 1991), a diferença entre a aproximação superior e a inferior é chamada de **região de fronteira** $B(X)$, representando a incerteza associada ao conjunto X e quantificando a imprecisão do conhecimento disponível. A qualidade da aproximação, por sua vez, é medida pela **acurácia da aproximação** $\alpha(X)$. Respectivamente:

$$[6] \quad B(X) = \overline{X} \setminus X = \{x \in U \mid x \in \overline{X} \wedge x \notin \underline{X}\},$$

$$[7] \quad \alpha(X) = \frac{|\underline{X}|}{|\overline{X}|}$$

Os fundamentos teóricos da *Rough Set Theory* (RST) se estendem aos aspectos topológicos, pois as uniões de classes elementares – como aproximações inferiores e superiores – formam uma estrutura que pode ser interpretada como uma álgebra booleana (PAWLAK, 1982, 1991). Nesse sistema, as operações de união, interseção e complemento resultam em

novos conjuntos que preservam a mesma estrutura, demonstrando as propriedades fundamentais de uma álgebra booleana (WHITESITT, 1995).

Além disso, os aspectos topológicos referem-se às propriedades dos conjuntos que são preservadas em transformações contínuas, tais como conectividade ou compacticidade (MUNKRES, 2000). Dessa forma, é possível interpretar as operações de aproximação na RST como os operadores de interior e fechamento em um espaço topológico cuja base é formada por essas classes elementares (PAWLAK, 1982, 1991).

Ademais, a capacidade de reduzir a dimensionalidade dos dados – por meio da identificação de atributos redundantes utilizando os conceitos de *reducts* e *core* – demonstra o poder dessa teoria para lidar com conjuntos de dados complexos, preservando a informação essencial para a classificação e a tomada de decisão. Em *Rough Set Theory*, os *reducts* são subconjuntos mínimos dos atributos que mantêm a mesma capacidade de classificação do conjunto original, ou seja, são as combinações mais simples de atributos que não perdem informações relevantes para a decisão. Já o *core* é a interseção de todos os *reducts*, representando os atributos indispensáveis que aparecem em todas as possíveis reduções (PAWLAK, 1991).

No campo das aplicações, a RST tem sido empregada com sucesso em problemas de classificação e tomada de decisão, destacando-se a extração de regras decisórias a partir de tabelas de decisão (PAWLAK, 1991). Um exemplo notório é o sistema “*Learning from Examples based on Rough Sets*”, que induz regras a partir de dados tabulados e permite tanto a classificação exata – baseada na aproximação inferior – quanto a classificação aproximada – fundamentada na aproximação superior (GRZYMALA-BUSSE, 1992). Essa abordagem, que já se consolidava na época de sua criação, foi aplicada em áreas tão diversas quanto medicina, pesquisa de mercado, processamento de imagem e análise de vibrações, assim como outras áreas (PAWLAK et al., 1995).

Entretanto, a estrutura clássica da RST, alicerçada em relações de equivalência, enfrenta limitações quando aplicada a contextos de decisão multicritério, onde a incorporação de relações de preferência ou ordenação é fundamental (GRECO, MATARAZZO e SLOWINSKI, 2001). Em cenários onde os dados apresentam níveis graduais de similaridade a abordagem binária de pertencimento ou não da RST pode não capturar nuances essenciais, o que tem motivado propostas de aprimoramento que não serão aprofundadas nesta subseção,

mas que são discutidas mais à frente (ZIARKO, 1993) (GRECO et al., 2001) (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Essa evolução teórica conduziu, inicialmente, ao *Indiscernibility Rough Set Approach* (IRSA) (GRZYMALA-BUSSE, 1992), que posteriormente evoluiu para o *Dominance-based Rough Set Approach* (DRSA) (GRECO, MATARAZZO e SLOWINSKI, 2001) – metodologias que se mostraram particularmente adequadas para problemas de classificação nominal e ordinal, respectivamente (GRECO, MATARAZZO e SLOWINSKI, 2001). Mais recentemente, para atender a contextos ainda mais complexos, foram desenvolvidas as versões com consistência variável, denominadas VC-IRSA e VC-DRSA, as quais ajustam o grau de tolerância à inconsistência e ampliam as possibilidades de análise (GRECO et al., 2001) (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Além das extensões teóricas, a aplicação prática da RST enfrenta desafios computacionais, sobretudo na determinação dos *reductos* em grandes bases de dados. Diversos estudos propuseram algoritmos heurísticos e metaheurísticos para mitigar esse custo computacional, tornando o processo de extração de conhecimento mais eficiente (PAWLAK et al., 1995) (KOMOROWSKI, POLKOWSKI e SKOWRON, 1999).

Para superar essas limitações, foram sugeridos caminhos como o desenvolvimento de algoritmos mais eficientes para a identificação de *reducts* e a extensão dos fundamentos teóricos da RST para incluir mecanismos que integrem relações de preferência e ordenação, ampliando assim a análise em contextos multicritério. Tais propostas derivam dos desdobramentos teóricos originais, que combinaram os fundamentos da RST com novas perspectivas na análise de dados, e continuam a inspirar pesquisas na área (SKOWRON, 2007).

Adicionalmente, a integração da RST com outras técnicas tem potencializado sua aplicabilidade. Por exemplo, a combinação da RST com métodos de lógica *fuzzy*, redes neurais e algoritmos evolutivos têm sido explorada para melhorar a extração de regras e a *precision* na classificação de dados, ampliando o espectro de problemas que podem ser abordados por sistemas inteligentes (SKOWRON e DUTTA, 2018).

Embora desafios persistam em ambientes de decisão multicritério, os fundamentos históricos, matemáticos e topológicos que embasaram a RST – desde seus primórdios até os estudos que expandiram seus horizontes – a consolidam como uma ferramenta

indispensável no arsenal das técnicas de análise de dados, inteligência artificial e descoberta de conhecimento (PAWLAK et al., 1995).

2.2. IRSA

A partir dos fundamentos teóricos estabelecidos pela *Rough Set Theory* (RST), que descreve a relação de indiscernibilidade, as aproximações inferior e superior, a região de fronteira e a acurácia da aproximação (PAWLAK, 1982, 1991), o conhecimento evoluiu para aplicações práticas na indução de regras, culminando no desenvolvimento do sistema LERS (*Learning from Examples based on Rough Sets*) – e, posteriormente, o que seria o IRSA (*Indiscernibility Rough Set Approach*) – por Grzymala-Busse no seu artigo “*LERS - A System for Learning from Examples Based on Rough Sets*” de 1992, e refinado em 1997 no seu artigo “*A New Version of the Rule Induction System LERS*”.

Enquanto que no RST, a inconsistência dos dados é tratada por meio do cálculo das aproximações, o sistema LERS aplica esses princípios à indução de regras a partir de tabelas de decisão. Para cada conceito definido por pares **atributo-valor** como (d, w) – em que d é o atributo de decisão e w um de seus valores possíveis – LERS computa as aproximações inferior e superior, gerando regras certas (derivadas da aproximação inferior) e regras possíveis (derivadas da aproximação superior). Além disso, o sistema associa medidas quantitativas ao conhecimento extraído, permitindo avaliar o desempenho das regras induzidas e refletindo, de forma mensurável, o grau de *precision* e confiabilidade obtido no processo de classificação:

$$[8] \quad q_{\underline{X}} = \frac{|\underline{X}|}{|U|}$$

$$[9] \quad q_{\overline{X}} = \frac{|\overline{X}|}{|U|}$$

que representam, respectivamente, a **qualidade da aproximação inferior** e a **qualidade da aproximação superior** dos exemplos ao conjunto X (GRZYMALA-BUSSE, 1997).

Inicialmente, o LERS era aplicado apenas a atributos categóricos; contudo, a evolução dos métodos permitiu que o sistema fosse aprimorado para lidar também com atributos numéricos por meio de discretização, ampliando assim sua aplicabilidade prática. Em seus primórdios, o LERS utilizava o algoritmo LEM (*Learning from Examples Module*), um método global que busca subconjuntos mínimos de atributos – os chamados *reducts* – capazes de preservar o poder discriminante do conceito. Entretanto, a necessidade de regras mais

simples e interpretáveis motivou o desenvolvimento do algoritmo LEM2, que adota uma estratégia de construção de coberturas locais (GRZYMALA-BUSSE, 1992, 1997).

Neste método, para cada conjunto E_T – associado à aproximação inferior ou superior de um conceito –, um conjunto mínimo T de condições que devem ser atendidas se, para qualquer subconjunto $T' \subset T$, $U[T']$ for diferente de E_T . Cada condição é definida por um par atributo–valor $t = (a, v)$ – em que a é um atributo de condição e v um de seus valores possíveis –, e o agrupamento de exemplos com o mesmo valor para esse atributo é denotado por $[t]$ (GRZYMALA-BUSSE, 1992, 1997):

$$[10] \quad U[T] = \bigcup_{t \in T} [t] = E_T$$

A seleção dos pares de atributo-valor no algoritmo LEM2 segue uma hierarquia de critérios: primeiramente, priorizam-se os atributos com maior importância; em seguida, avalia-se o tamanho da interseção entre $[t]$ e um conjunto G de exemplos; e, por fim, utiliza-se a cardinalidade $|[t]|$ como critério de desempate. Essa estratégia contribui para a obtenção de regras mais simples, interpretáveis e eficientes na classificação (GRZYMALA-BUSSE, 1992, 1997).

Em 1992, Grzymala-Busse demonstrou a eficácia do LERS na indução de regras classificatórias em dados inconsistentes, gerando regras certas e possíveis. Em 1997, o sistema foi aprimorado para lidar com atributos numéricos por meio de discretização e incorporar ferramentas de validação – como *leaving-one-out*, *ten-fold cross validation* e *hold-out* (GRZYMALA-BUSSE, 1992, 1997). Assim, a evolução dos métodos partiu dos fundamentos precisos da RST para sistemas como o LERS e culminou no IRSA, que utiliza as aproximações para identificar regras classificatórias e integra mecanismos multicritério, permitindo um tratamento refinado dos dados com níveis graduais de similaridade e imprecisão (PAWLAK et al., 1995; SKOWRON, 2007; ZIARKO, 1993).

2.3. DRSA

Na evolução das técnicas de análise de dados multicritério, o *Dominance-based Rough Set Approach* (DRSA) emerge como uma proposta inovadora para lidar com problemas decisórios onde a ordenação dos critérios é determinante (GRECO, MATARAZZO e SLOWINSKI, 2001) (BLASZCZYNSKI et al., 2022). Em contraste com a abordagem clássica da RST, que utiliza a relação de indiscernibilidade para particionar o universo dos dados em

classes elementares (PAWLAK, 1982, 1991) (GRZYMALA-BUSSE, 1992, 1997), o DRSA fundamenta-se na relação de dominância. Essa mudança conceitual permite que a análise incorpore diretamente a noção de preferência, essencial em contextos em que os critérios possuem uma ordem intrínseca, como em problemas de avaliação de desempenho, classificação de riscos ou ordenação de alternativas (GRECO, MATARAZZO e SLOWINSKI, 2001).

Diferentemente dos atributos condicionais empregados na RST clássica – que apenas descrevem os objetos sem necessariamente refletir uma hierarquia ou ordenação de importância – em DRSA o termo “critério” assume um papel central. Cada critério $c \in P$ está associado a uma função de avaliação f_c , que mapeia os objetos para uma escala ordenada representando as preferências do decisor. Adicionalmente, os critérios podem ser classificados em dois tipos: de ganho, para os quais valores mais altos são preferíveis, e de custo, em que valores menores são desejáveis. Essa ordenação intrínseca possibilita definir a relação de dominância de forma que a comparação entre objetos incorpore não apenas a verificação de igualdade ou diferença, mas também a intensidade da preferência (GRECO, MATARAZZO e SLOWINSKI, 2001).

Portanto, define-se que um objeto x domina um objeto y se, para cada critério c , a avaliação de x não for inferior à avaliação de y e, adicionalmente, existir ao menos um critério para o qual a avaliação de x seja estritamente maior. Formalmente, expressa-se essa relação como:

$$[11] \quad xD_p y \Leftrightarrow \bigwedge_{c \in P} f_c(x) \geq f_c(y)$$

$$[12] \quad xD^s_p y \Leftrightarrow (xD_p y) \bigwedge (\exists c \in P: f_c(x) > f_c(y))$$

essa formulação, é capaz de captar a estrutura ordinal dos dados, pois não se restringe apenas à igualdade ou à diferença entre atributos, mas incorpora também a qualidade relativa do desempenho dos objetos. Dessa forma, o DRSA se torna especialmente útil em contextos decisórios, onde é fundamental distinguir entre alternativas que, embora apresentem diferenças mínimas, podem possuir níveis de desempenho significativamente distintos quando avaliados em múltiplos critérios (GRECO, MATARAZZO e SLOWINSKI, 2001).

A partir dessa relação, são construídos os cones de dominância, que são conjuntos de objetos que sintetizam a ordem estabelecida. O cone de dominância positivo de um objeto x , denominado $D_p^+(x)$, agrupa todos os objetos que x domina, isto é, todos aqueles cuja

avaliação é inferior ou igual à de x em todos os critérios e, em pelo menos um critério, estritamente inferior. Por outro lado, o cone de dominância negativo de x , chamado $D_P^-(x)$, reúne todos os objetos que dominam x , ou seja, aqueles cuja avaliação é superior ou igual à de x em todos os critérios, com pelo menos um critério onde seja estritamente superior. Esses cones são essenciais para que o DRSA possa capturar não só a similaridade, mas também a intensidade relativa das avaliações, refletindo a estrutura ordinal dos dados (GRECO, MATARAZZO e SLOWINSKI, 2001).

Com a substituição da relação de indiscernibilidade pela relação de dominância, os conceitos de aproximação inferior e superior também sofrem uma redefinição para melhor refletir os cenários decisórios. Considerando um universo de objetos U e uma classe de decisão Cl_t , define-se a **aproximação inferior no DRSA** como o conjunto dos objetos cuja dominância garante que todos os objetos que os dominam pertençam definitivamente à classe Cl_t . Em termos formais:

$$[13] \quad P^-(Cl_t) = \{x \in U : \forall_y | y D_P x, y \in Cl_t\}$$

isso significa que, se um objeto x possui a propriedade de que todos os objetos que o dominam estão na classe Cl_t , então x pode ser classificado com certeza como pertencente a Cl_t (GRECO, MATARAZZO e SLOWINSKI, 2001).

Por outro lado, a **aproximação superior no DRSA** é definida como o conjunto dos objetos para os quais existe ao menos um objeto dominado que pertence à classe Cl_t . Formalmente, escreve-se:

$$[14] \quad P^+(Cl_t) = \{x \in U : \exists_y | x D_P y \wedge y \in Cl_t\}$$

essa definição permite que o DRSA capture a incerteza e a ambiguidade inerentes aos dados, ao identificar objetos que, embora possam não ser classificados com total certeza, possuem alguma associação com a classe Cl_t por meio da dominância. Assim, os conceitos de aproximação inferior e superior não apenas preservam a lógica fundamental dos *rough sets*, mas também ampliam sua aplicabilidade para contextos em que a ordem de preferência é essencial (GRECO, MATARAZZO e SLOWINSKI, 2001).

Uma das contribuições mais significativas do DRSA é sua capacidade de gerar regras decisórias que respeitam a lógica da dominância. Ao invés de basear-se apenas na similaridade entre atributos, o DRSA permite extrair regras na forma de declarações do tipo:

$$[15] \quad \bigwedge_{i=1}^k (c_j \in V_i) \Rightarrow d = Cl_t$$

onde cada condição $c_i \in V_i$ é escolhida de modo a preservar a monotonicidade dos dados. Em outras palavras, se um objeto satisfaz uma determinada condição, todos os objetos dominados por ele também devem satisfazer essa condição. Essa característica é fundamental para manter a coerência com a estrutura ordinal dos dados e reflete a ideia de que, ao melhorar a avaliação de um critério, não se deve prejudicar a classificação global (GRECO, MATARAZZO e SLOWINSKI, 2001).

Ademais, o DRSA amplia o conceito tradicional de *reducts* – que, na Teoria do RTS clássica, correspondem a subconjuntos mínimos de atributos capazes de preservar a mesma classificação do conjunto original – ao identificar subconjuntos mínimos de critérios que mantêm a capacidade discriminatória dos objetos. Essa redução é crucial para diminuir a complexidade do modelo decisório, permitindo a formação de regras robustas e interpretáveis, mesmo em conjuntos de dados com alta dimensionalidade (GRECO, MATARAZZO e SLOWINSKI, 2001).

A abordagem DRSA, ao utilizar a relação de dominância, oferece uma representação mais fiel dos problemas decisórios multicritério, pois não só discrimina objetos com atributos idênticos, mas também considera a qualidade relativa dos desempenhos, refletindo a estrutura ordinal dos dados (BLASZCZYNSKI et al., 2022). Essa metodologia é valiosa para áreas como análise de risco, avaliação de desempenho, seleção de investimentos e classificação de alternativas (CHAKHAR e SAAD, 2012; SZELAZG, GRECO e SLOWINSKI, 2014); PAWLAK et al., 1995; BLASZCZYNSKI et al., 2022), especialmente porque permite a extração de regras em linguagem natural – "se ... então ..." – que facilitam a interpretação e a justificativa das escolhas (SLOWINSKI, GRECO e MATARAZZO, 2015). Além disso, a identificação de *reducts* e a determinação do *core* reduzem a dimensionalidade do problema, otimizando o processo decisório e melhorando a transparência dos resultados (GRECO et al., 2000).

O algoritmo DomLEM operacionaliza esses conceitos de maneira iterativa e adaptativa, refinando continuamente o conjunto de objetos a serem classificados ao remover exemplos já cobertos, o que o torna robusto para ambientes de dados complexos. Dessa forma, tanto do ponto de vista teórico quanto prático, o DRSA e o DomLEM consolidam-se como metodologias indispensáveis para a modelagem e resolução de problemas decisórios

complexos, alinhando os sistemas de apoio à decisão com a lógica natural dos critérios de avaliação (GRECO et al., 2000; BLASZCZYNSKI et al., 2022).

2.4. DomLEM

O *Dominance-based Learning from Examples Module* (DomLEM) é um algoritmo de indução de regras projetado para aplicações multicritério que gera um conjunto mínimo de regras decisórias alinhadas à lógica de dominância e que atendem aos requisitos de confiabilidade e interpretabilidade. Desenvolvido no contexto da DRSA, o DomLEM adapta o LEM2 – originalmente criado para a *Rough Set Theory* baseada na indiscernibilidade – (GRZYMALA-BUSSE, 1992, 1997) para lidar especificamente com problemas de decisão multicritério nos quais a ordenação dos critérios é fundamental (GRECO et al., 2000). Segundo Greco et al., (2000), funcionamento do DomLEM pode ser descrito em etapas detalhadas:

- **Entrada e Preparação dos Dados:** O algoritmo recebe como entrada um conjunto de objetos U descritos por atributos condicionais C e um atributo decisório d . Junto a isso, são fornecidas as funções de avaliação f_c que estabelecem a relação de dominância DP para um conjunto de critérios $P \subseteq C$ e as uniões de classes ascendentes e descendentes.
- **Definição do Conjunto Alvo:** É definido um subconjunto $T \subseteq U$ que contém os objetos pertencentes à classe decisória de interesse Cl_t . Esse conjunto serve como base para a indução das regras que deverão cobrir os exemplos positivos.
- **Seleção Iterativa de Pivôs e Construção de Condições:** O DomLEM adota uma estratégia iterativa na qual, a cada iteração, um objeto *pivot* é selecionado a partir de T . Em seguida, o algoritmo constrói um conjunto inicial de condições. Nesse processo, busca-se adicionar, de forma iterativa, condições que maximizem a cobertura dos objetos positivos – aqueles que pertencem a Cl_t – e minimizem a inclusão de objetos negativos. Para orientar essa seleção, utiliza-se uma métrica que relaciona a cobertura dos objetos positivos à cobertura total dos objetos avaliados, representada pela razão:

$$[16] \quad \Delta = \frac{|P^-(Cl_t)|}{|U|}$$

- **Verificação de Consistência e Nível de Confiança:** Durante a construção do conjunto de condições (denotado por T'), é feita uma verificação para assegurar que o conjunto acumulado esteja contido na aproximação inferior da classe Cl_t ou atinja um nível de confiança pré-estabelecido. Essa etapa é crucial para garantir que as regras geradas não

apenas sejam representativas dos dados, mas também possuam respaldo empírico – ou seja, existam objetos que validem a regra.

- **Geração e Refinamento das Regras Decisórias:** Uma vez que o conjunto de condições satisfaz os critérios de consistência, o algoritmo gera a regra no formato “se $c_1 \in V_1$ e ... e $c_k \in V_k$, então $d = Cl_t$ ”. Posteriormente, é realizada uma etapa de poda que visa remover condições redundantes. Cada condição é testada de modo que, se sua exclusão não comprometer significativamente a acurácia ou a cobertura da regra, ela é removida. Esse refinamento resulta em regras mais concisas e facilmente interpretáveis.
- **Atualização e Iteração do Conjunto de Objetos:** Após a geração de uma regra robusta – isto é, uma regra que possui ao menos um objeto base (um exemplo real que satisfaz exatamente as condições da regra) – os objetos já classificados de forma satisfatória são removidos do conjunto T . O processo se repete iterativamente até que todos os objetos relevantes sejam cobertos ou até que não seja possível obter melhorias adicionais.

O DomLEM integra a relação de dominância com as etapas de indução e refinamento das regras, refletindo a convergência entre os fundamentos teóricos da DRSA e sua aplicação prática. Essa abordagem garante um modelo rigoroso e adaptável a ambientes decisórios complexos, tendo sido aplicado com sucesso em diversas áreas, como avaliação de desempenho, seleção de alternativas, diagnóstico médico, análise de sustentabilidade empresarial e satisfação do cliente (GRECO et al., 2000)..

Apesar de sua eficácia, o algoritmo enfrenta limitações, como alta complexidade computacional para grandes bases de dados, sensibilidade a ruídos devido à rigidez das condições de inclusão, pressuposição de relações de preferência monotônicas e restrição no tratamento de atributos nominais. Para contornar esses desafios, foram propostas extensões como o VC-DomLEM, que, através de abordagens de consistência variável (como VC-IRSA e VC-DRSA) e conceitos oriundos do VPRS, torna o modelo mais robusto e capaz de gerar regras mais gerais, ampliando as ferramentas de análise baseadas em DRSA (GRECO et al., 2000).

2.5. VC-IRSA E VC-DRSA

A *Rough Set Theory* de Pawlak (1982, 1991) utiliza aproximações inferiores e superiores para tratar imprecisão nos dados, mas suas definições clássicas – baseadas em indiscernibilidade ou dominância – exigem inclusão total de um objeto em seu grânulo de conhecimento ou cone de dominância, gerando aproximações inferiores muito restritas ou vazias em bases ruidosas (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009). Para contornar

isso, o conceito de consistência variável permite relaxar esse requisito, incluindo na aproximação inferior objetos com forte – mas não absoluta – relação ao conjunto (GRECO et al., 2001; BŁASZCZYŃSKI et al., 2006). Essa flexibilidade equilibra certeza e cobertura, tornando os modelos mais robustos para dados complexos (SZELAG, GRECO e SLOWINSKI, 2014).

Para superar essa limitação, foram propostos modelos probabilísticos e de consistência variável – denominados *Variable Consistency Indiscernibility-based Rough Set Approach* (VC-IRSA) e *Variable Consistency Dominance-based Rough Set Approach* (VC-DRSA) – que introduzem uma medida de consistência de objetos θ_X para quantificar a evidência de pertencimento em aproximações inferiores, sujeita a um limiar de consistência de objetos, denotado por θ_X , onde $0 \leq \theta_X \leq A_X$. Esse parâmetro controla o nível mínimo e máximo – este dentro da escala do critério – de evidência exigido para que um objeto seja considerado como pertencente ao conjunto aproximado.

Dessa forma, os modelos VC IRSA e VC DRSA oferecem um equilíbrio entre rigor e flexibilidade, possibilitando uma modelagem mais robusta de dados com inconsistências ou incertezas (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009). Essa abordagem reconhece que, em aplicações práticas, pode ser vantajoso relaxar os critérios estritos de inclusão para capturar objetos que possuem uma forte, embora não totalmente certa, relação com o conjunto aproximado (BŁASZCZYŃSKI et al., 2006). A introdução de limiares de consistência permite que o modelo tolere um certo grau de imprecisão, tornando-o mais adaptável a dados ruidosos ou com sobreposição de classes (KARAMI, GORSEVSKI e JANKOWSKI, 2014). A avaliação da consistência, seja através de medidas relativas ou absolutas, oferece uma maneira quantitativa de determinar a evidência de pertinência, permitindo uma construção de aproximações inferiores mais informativas e úteis para a indução de regras (LI et al., 2020)

Por questão de simplicidade, será definido que X representará tanto objetos pertencentes à uma classe X_i , no contexto de VC-IRSA, como em uma união de classes ascendentes X_i^{\geq} e descendentes X_i^{\leq} , como de custo, onde menor é melhor. Sendo assim a aproximação inferior é definida como:

$$[17] \quad \underline{X} = \{y \in U: \theta_X(y) \propto \theta_X\}$$

Onde \propto denota \geq para medidas de ganho e \leq para medidas de custo (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

A flexibilização introduzida por θ_X permite que o modelo seja ajustado conforme a natureza dos dados. Por exemplo, em contextos de diagnóstico médico, onde a presença de ruído ou casos atípicos é comum, um valor de θ_X menor pode ampliar a cobertura dos conjuntos inferiores, permitindo a inclusão de pacientes com sintomas majoritariamente compatíveis com uma determinada doença, mesmo que haja exceções (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Segundo Błaszczński (2009, 2011), a flexibilização da condição de inclusão nos modelos de consistência variável requer a introdução de medidas que quantifiquem a evidência de que um objeto y pertence a um conjunto X . Tais medidas são geralmente classificadas em tipos de ganho e custo. Antes de introduzirmos as medidas propriamente ditas, é fundamental definir as propriedades de monotonicidade que elas devem satisfazer para que as aproximações inferiores resultantes preservem a estrutura desejada dos *rough sets*. Inspirados nos fundamentos da RST, essas propriedades podem ser organizadas em diferentes dimensões:

- (m1) Monotonicidade em relação ao conjunto de atributos: Ao aumentar o conjunto de atributos, a evidência de que um objeto pertence a X não deve diminuir. Em termos formais, para todos $P \subseteq P' \subseteq C$ e para um objeto y respectivamente em medidas de ganho e custo, espera-se que:

$$[18] \quad \theta_X^P(y) \leq \theta_X^{P'}(y)$$

$$[19] \quad \theta_X^P(y) \geq \theta_X^{P'}(y)$$

Essa propriedade garante que a adição de atributos – que resulta em uma descrição mais detalhada – não reduza a evidência de associação dos objetos a X (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

- (m2) Monotonicidade em relação ao conjunto de objetos: Quando o conjunto X é expandido com a inclusão de novos objetos, a evidência para a associação dos objetos já presentes não deve piorar. Formalmente, se $X \subseteq X' \subseteq U$, respectivamente para medidas de ganho e custo deve ocorrer que:

$$[20] \quad \theta_X(y) \leq \theta_{X'}(y)$$

$$[21] \quad \theta_X(y) \geq \theta_{X'}(y)$$

Essa propriedade assegura que a expansão de X com novos exemplos não comprometa a consistência dos objetos previamente classificados (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

- (m3) Monotonicidade em relação à união de classes (específica para DRSA): Em problemas com classes ordenadas, quando se tem uma união de classes X_i que está contida em uma união mais ampla X_j , espera-se que a evidência de associação para X_i seja menor que X_j em critérios de ganho e maior que em critérios de custo, respectivamente:

$$[22] \quad \theta_{X_i^*}(y) \leq \theta_{X_j^*}(y)$$

$$[23] \quad \theta_{X_i^*}(y) \geq \theta_{X_j^*}(y)$$

Onde “*” representa tanto “ \geq ” como “ \leq ”. Esta propriedade reflete a ideia de que conjuntos mais amplos devem apresentar uma evidência de associação mais forte do que conjuntos mais restritos (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

- (m4) Monotonicidade em relação à dominância (específica para DRSA): Quando a relação de dominância é considerada, se um objeto y_1 domina y_2 – denotado por $y_1 D_P y_2$ – então, para uma medida do tipo ganho, a evidência de associação de y_1 a X não deve ser inferior à de y_2 :

$$[24] \quad \forall_{y_1, y_2 \in U}: y_1 D y_2 \Rightarrow \theta_{X_i^*}(y_1) \geq \theta_{X_i^*}(y_2)$$

$$[25] \quad \forall_{y_1, y_2 \in U}: y_1 D y_2 \Rightarrow \theta_{X_i^*}(y_1) \leq \theta_{X_i^*}(y_2)$$

Essa propriedade garante que a hierarquia estabelecida pela dominância seja respeitada na quantificação da evidência de associação, mantendo a coerência dos modelos em contextos ordenados (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Para VC-RSA, define-se também as **regiões positiva, negativa e de fronteira** de X como:

$$[26] \quad POS(X) = \bigcup_{y \in \underline{X}} E(y)$$

$$[27] \quad NEG(X) = POS(\neg X) - POS(X)$$

$$[28] \quad BND(X) = (U - POS(X)) - NEG(X)$$

Em que região positiva $POS(X)$ representa todo o “espaço” que pode ser alcançado pelos elementos efetivamente em X . Em contrapartida, a região negativa $NEG(y)$ reúne os elementos que certamente não estão em X . Já a região de fronteira $BND(X)$ corresponde aos objetos que os quais não se pode ter certeza se estão em X ou não, representando a zona de incerteza na classificação (PAWLAK, 1982; BŁASZCZYŃSKI et al., 2011; SLOWINSKI, GRECO e MATARAZZO, 2015).

Pesquisas recentes introduziram medidas de consistência robustas que preservam propriedades de monotonicidade e regiões definidas (BŁASZCZYŃSKI, SLOWINSKI e SZELAG, 2009). Nos modelos VC-RSA, níveis variáveis de consistência aumentam a flexibilidade das aproximações inferiores e superiores, permitindo inconsistência controlada (BŁASZCZYŃSKI et al., 2006) e evitando limiares rígidos que geram aproximações excessivamente pequenas (LI et al., 2020). No multicriteria ranking, o VC-DRSA aborda inconsistências nas preferências dos decisores ao selecionar pares coerentes para induzir modelos de preferência (SZELAG, GRECO e SLOWINSKI, 2014). Em aplicações como análise da qualidade da água, o relaxamento do nível de consistência captura nuances dos dados ambientais (KARAMI, GORSEVSKI e JANKOWSKI, 2014).

- Para modelos baseados em custo, a melhor medida trata-se de ϵ^*X , definida para as uniões de classe ascendente e descendente, respectivamente:

$$[29] \quad \epsilon_{X_i^>}^*(y) = \max_{j \leq i} \epsilon_{X_j^>}(y)$$

$$[30] \quad \epsilon_{X_i^<}^*(y) = \max_{j \geq i} \epsilon_{X_j^<}(y)$$

$$[3'] \quad \text{dado } \epsilon_X(y) = \frac{|E(y) \cap \neg X|}{|\neg X|}$$

As quais corrigem a falta da propriedade ($m3$) presente na medida original ϵX , além de:

$$[32] \quad \epsilon'_X(y) = \frac{|E(y) \cap X|}{|E(y)|}$$

as quais garantem que todas as quatro propriedades de monotonicidade sejam satisfeitas no contexto do DRSA (BŁASZCZYŃSKI et al., 2011).

- Para modelos baseados em ganho, a melhor medida trata-se de $\mu'X$, definida para as uniões de classe ascendente e descendente, respectivamente:

$$[33] \quad \mu'_{X_i^>}(y) = \max_{z \in D^-(y) \cap X_i^>} \mu_{X_i^>}(z)$$

$$[34] \quad \mu'_{X_i^{\leq}}(y) = \max_{z \in D^-(y) \cap X_i^{\leq}} \mu_{X_i^{\leq}}(z)$$

$$[35] \quad \text{dado } \mu_X(y) = \frac{|E(y) \cap X|}{|E(y)|}$$

Essa medida expande a medida de *rough membership* μ_X , garantindo que as propriedades (m2), (m3) e (m4) sejam satisfeitas, embora ainda falte a propriedade (m1) (BŁASZCZYŃSKI et al., 2011).

Dessa forma, a escolha adequada de medidas de consistência é fundamental para tornar as aproximações inferiores em modelos *Rough Sets* de consistência variável mais robustas. Propostas como as de Błaszczynski et al. (2011) corrigem limitações anteriores, preservando propriedades essenciais e aumentando a confiabilidade na geração de regras decisórias. Em VC-DRSA, diferentes medidas refinam a seleção de objetos para aproximações inferiores segundo critérios de consistência (SZELAG, GRECO e SLOWINSKI, 2014), enquanto níveis quantitativos de consistência melhoram a tolerância a falhas e incorporam informações numéricas (LI et al., 2020). Essas definições são cruciais para a eficácia dos modelos VC-RSA em variadas aplicações.

As regras de decisão seguem o formato *se Φ então Ψ* , com Φ como premissa e Ψ como conclusão, e são induzidas para cobrir objetos nas aproximações inferiores (probabilísticas) de classes ou uniões de classes. Essa flexibilidade reconhece que dados reais frequentemente não permitem certeza absoluta na classificação (LI et al., 2020). Assim, considera-se a região positiva $POS(X)$, que inclui objetos com forte evidência de pertencimento, mesmo sem certeza completa (SZELAG, GRECO e SLOWINSKI, 2014). Cada regra deve ser suportada por ao menos um objeto de X e cobrir objetos de $POS(X)$. Níveis variáveis de consistência acomodam diferentes graus de incerteza, produzindo modelos mais robustos para dados complexos (FAN, LIAU e LIU, 2013). No VC-IRSA – onde não há ordenação natural entre atributos – cada regra pode ser expressa como:

$$[36] \quad \bigwedge_{k=1}^z [q_{ik} = t_{ik}] \Rightarrow y \in X_i$$

onde q_{ik} são atributos regulares, e t_{ik} é um valor de y em cada atributo. No entanto, quando os dados são ordenados ou há preferência em alguns critérios, como no VC-DRSA, a sintaxe da regra reflete a monotonicidade para a união ascendente e descendente, respectivamente:

$$[37] \quad \left(\bigwedge_{j=1}^p [q_{ij} \geq t_{ij}] \wedge \bigwedge_{k=p+1}^z [q_{ik} = t_{ik}] \right) \Rightarrow y \in X_i^{\geq},$$

$$[38] \quad \left(\bigwedge_{j=1}^p [q_{ij} \leq t_{ij}] \wedge \bigwedge_{k=p+1}^z [q_{ik} = t_{ik}] \right) \Rightarrow y \in X_i^{\leq}$$

Em que q_i representa os atributos condicionais – seja como critérios em q_{ij} ou como atributos regulares em q_{ik} . Em paralelo, a distinção entre critério de ganho e critério de custo faz com que “ $q_{ij}(y) \leq t_{ij}$ ” signifique “ y não é pior que t_{ij} ” em casos de critério de ganho, enquanto para critério de custo esse mesmo símbolo pode significar “ y não é melhor que t_{ij} ”. Assim, há duas camadas de interpretação: (1) uniões ascendentes ou descendentes – X_i^{\geq} e X_i^{\leq} , respectivamente – e (2) a natureza do critério (ganho ou custo), que define a forma exata de ler “ \geq ” ou “ \leq ” – os quais representam fraca preferência e fraca preferência inversa, respectivamente (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Portanto, as regras devem ser curtas, para serem fáceis de entender e menos propensas à sobreajuste, e acuradas, para classificar corretamente novos objetos. Dessa forma, a formulação possibilita que o decisor não só compreenda as condições associadas à decisão, mas também avalie o grau de confiança – controlado por $\hat{\theta}_X$ – atribuído a cada regra (BŁASZCZYŃSKI et al., 2011).

Para tanto, podem ser utilizadas diversas abordagens para avaliar a consistência de uma regra, denotada por r_X , que propõe a atribuição de objetos ao conjunto X . Seja $\Phi(r_X)$ o conjunto de condições que satisfazem a decisão $\Psi(r_X)$, e seja $|\Phi(r_X)|$ o conjunto de objetos que atendem a essas condições (FAN, LIAU e LIU, 2013). Além disso, definimos $\hat{\theta}_X$ como a medida de consistência da regra no contexto das aproximações inferiores, e R_X como o conjunto de todas as regras que atribuem um objeto a X (KARAMI, GORSEVSKI e JANKOWSKI, 2014). Segundo Błaszczński et al. (2009), são consideradas três medidas de consistência para as regras:

1. Consistência ϵ de r_X relacionada à medida de consistência de objetos de custo ϵ_X :

$$[39] \quad \hat{\epsilon}_X(r_X) = \frac{|\Phi(r_X)| \cap \neg \underline{X}|}{|\neg \underline{X}|}, \quad \text{sujeita ao limiar } \hat{\theta}_X = \frac{|\neg X|}{|\neg \underline{X}|}$$

Essa medida avalia, de forma relativa, quantos objetos fora de X ($\neg X$) são indevidamente cobertos pela regra;

2. Consistência ϵ' de r_X : relacionada à medida de consistência de objetos de custo ϵ'_X :

$$[40] \quad \hat{\epsilon}'_X(r_X) = \frac{||\Phi(r_X)|| \cap \neg X|}{|X|}, \quad \text{sujeita ao limiar } \hat{\theta}_X = \frac{|X|}{|X|}$$

Essa medida é similar à anterior, porém a normalização é feita em relação ao tamanho de X ;

3. Consistência μ de r_X relacionada à medida de pertinência aproximada do tipo ganho μ_X :

$$[41] \quad \hat{\mu}_X(r_X) = \frac{||\Phi(r_X)|| \cap X|}{||\Phi(r_X)||}, \quad \text{sujeita ao limiar } \hat{\theta}_X = \theta_X$$

Essa medida verifica a proporção de objetos cobertos pela regra r_X que de fato pertencem a X .

A primeira medida avalia, em relação ao tamanho de $\neg X$, quantos objetos fora de X foram indevidamente cobertos pela regra; a segunda faz o mesmo, mas normaliza pelo tamanho de X ; e a terceira calcula a proporção de objetos cobertos pela regra r_X que realmente pertencem a X . Diferentemente das definições originais para objetos, essas medidas penalizam a cobertura de itens fora da região positiva $POS(X)$, sobretudo os da região negativa $NEG(X)$, garantindo *precision* e reforçando propriedades essenciais de discriminação, minimalidade, completude e não-redundância na indução de regras (BŁASZCZYŃSKI et al., 2011; SZELAG, GRECO e SLOWINSKI, 2014). Assim, contribuem para modelos de decisão mais robustos, interpretáveis e transparentes (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009).

Os modelos VC-IRSA e VC-DRSA geram regras de decisão explícitas, caracterizando-se como métodos *glass-box* cujo processo de classificação é totalmente rastreável e interpretável pelo decisor (BŁASZCZYŃSKI et al., 2011). Essa transparência permite compreender a lógica das decisões ao examinar as condições que conduzem a cada classificação, ao contrário de abordagens *black-box* – como certos algoritmos de aprendizado profundo – cuja lógica interna é opaca (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009). Em aplicações que exigem alta interpretabilidade, como diagnósticos médicos ou auditorias

financeiras, a clareza das regras facilita a validação e aumenta a confiança dos especialistas no conhecimento extraído (SZELAG, GRECO e SLOWINSKI, 2014; LI et al., 2020).

Em conclusão, os modelos VC-IRSA e VC-DRSA aprimoram os *Rough Sets* clássicos ao introduzir o parâmetro de consistência λ , aumentando a tolerância a ruído e inconsistências comuns em dados reais (LI et al., 2020; FAN, LIAU e LIU, 2013). Suas medidas de consistência, que respeitam propriedades de monotonicidade, permitem aproximações inferiores mais flexíveis e informativas, capturando relações fortes mesmo sem completude total (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2009; SZELAG, GRECO e SLOWINSKI, 2014). Essa combinação de teoria robusta e indução prática de regras expande a aplicabilidade dos *Rough Sets*, gerando modelos decisórios robustos, interpretáveis e adaptáveis a contextos desafiadores, como análise de qualidade da água e diagnóstico industrial (KARAMI, GORSEVSKI e JANKOWSKI, 2014; LEZANSKI e PILACINSKA, 2016).

2.6. Avanços em *Rough Sets* e Suas Aplicações

A evolução dos *Rough Sets* tem sido marcada por inovações que flexibilizam a definição clássica da aproximação inferior, permitindo a inclusão de objetos com diferentes graus de certeza (SHI et al., 2024). Uma das primeiras abordagens a propor essa flexibilidade foi a *Variable Precision Rough Set* (VPRS), também conhecida como VP-RSA ou VP-IRSA, conforme apresentado por Ziarko (1993). Essa proposta representa uma extensão significativa do RST tradicional, abrindo caminho para desenvolvimentos posteriores, como o VC-DRSA, e promovendo uma análise mais adaptativa dos dados.

As principais características do VPRS incluem o relaxamento da certeza, que difere do método clássico ao admitir que a aproximação inferior incorpore objetos com uma relação forte – embora não perfeita – com os *rough sets*, controlada por um nível de certeza predefinido que permite a presença de um número limitado de contraexemplos (ZIARKO, 1993; GONG e LI, 2023). Além disso, a metodologia emprega uma função de pertinência aproximada para testar e determinar a inclusão desses objetos, quantificando o grau de associação aos conjuntos em análise (KUSUNOKI et al., 2021). Essa abordagem compartilha uma intuição semelhante à do *Variable Consistency Dominance-based Rough Set Approach* (VC-DRSA), pois ambas introduzem parâmetros de controle que flexibilizam o critério de inclusão e ampliam as possibilidades de modelagem dos dados (BŁASZCZYŃSKI et al., 2006).

Além do VPRS, outras propostas visaram ampliar os fundamentos do RST clássico. Entre essas, destacam-se as abordagens probabilísticas, que utilizam o fator de Bayes ou a confirmação de medições para possibilitar a inclusão de objetos na aproximação inferior com base em evidências probabilísticas, ampliando o conjunto de elementos considerados quando há indícios suficientemente fortes de pertencimento (BŁASZCZYŃSKI et al., 2006; BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011). Paralelamente, os *Parameterized Rough Sets* (PRS) e o VC-IRSA introduzem parâmetros adicionais para ajustar a inclusão de objetos, contribuindo para modelos mais flexíveis e adaptáveis às características específicas dos dados (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011).

O conceito de região positiva, fundamental para a indução de regras de decisão no VC-DRSA, encontra correspondência com a ideia de aproximação inferior estendida no VPRS. Tanto Ślęzak, Greco e Slowinski (2014) quanto Ziarko (1993) definiram essa região – tanto no contexto do VPRS quanto dos *Rough Sets* Bayesianos – como a base para a indução de regras, demonstrando a integração teórica entre essas abordagens (BŁASZCZYŃSKI et al., 2006).

A integração do DRSA e de suas variações com outras técnicas de inteligência artificial tem sido fundamental para superar limitações individuais na modelagem decisória (SHI et al., 2024). A combinação do DRSA com lógica *fuzzy*, redes neurais, algoritmos genéticos e métodos estatísticos permite capturar a gradualidade e a incerteza dos dados, resultando em modelos mais robustos e interpretáveis (SLOWINSKI, GRECO e MATARAZZO, 2015; PAWLAK e SKOWRON, 2007; KOMOROWSKI, POLKOWSKI e SKOWRON, 1999).

Além disso, algoritmos clássicos como o DOMLEM e o VC-DomLEM têm passado por contínuos processos de desenvolvimento e refinamento. Esses aprimoramentos visam não apenas otimizar o desempenho computacional, mas também adaptar os algoritmos a dados com características específicas – como o tratamento de valores faltantes por meio de técnicas de imputação – e explorar novas medidas de consistência para a inclusão de objetos (TOPAL, GULER e UCAN, 2024; BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011).

As abordagens baseadas no DRSA e suas extensões têm sido aplicadas em diversos domínios, demonstrando grande versatilidade e eficácia na análise de dados complexos. Entre as aplicações, destacam-se o apoio à decisão multicritério (MCDA), problemas de classificação ordinal, ranking, avaliação de sustentabilidade empresarial,

prognósticos em contextos médicos – como encaminhamento para cuidados paliativos e diagnóstico de meningite – e diagnósticos industriais, como em processos de retificação (DO COUTO e RANGEL, 2024; GIL-HERRERA et al. 2015; GOWIN et al. 2007; LEZANSKI e PILACINSKA, 2006; SLOWINSKI, GRECO e MATARAZZO, 2015; TOPAL, GULER e UCAN, 2024).

Apesar dos avanços, permanecem desafios significativos, como a escalabilidade para conjuntos de dados massivos, a seleção e validação dos parâmetros de consistência, a manutenção da interpretabilidade em modelos híbridos complexos e a consolidação de uma base teórica abrangente para todas as variações do DRSA. Tais questões continuam a representar áreas promissoras para futuras pesquisas, exigindo o desenvolvimento de soluções inovadoras que garantam a robustez e a aplicabilidade desses métodos (KOMOROWSKI, POLKOWSKI e SKOWRON, 1999; GRECO et al., 2000; ZIARKO, 1993; BŁASZCZYŃSKI et al., 2006; GIL-HERRERA et al. 2015).

O VPRS (ou VP-RSA/VP-IRSA) foi pioneiro ao flexibilizar a definição da aproximação inferior, permitindo a inclusão controlada de contraexemplos e estabelecendo uma conexão teórica com o VC-DRSA e outras abordagens probabilísticas (KUSUNOKI et al., 2021). A hibridização do DRSA com outras técnicas de inteligência artificial tem ampliado o leque de aplicações, proporcionando modelos decisórios mais robustos e precisos (SHI et al., 2024). Contudo, desafios relacionados à escalabilidade, à escolha dos parâmetros e à manutenção da interpretabilidade continuam impulsionando a pesquisa nesta área (OGANGE, 2023).

Rough Sets evoluiu significativamente desde as propostas iniciais de Pawlak (1982) que introduziram o conceito de aproximações inferiores e superiores, passando pelo desenvolvimento do IRSA e DRSA, os quais incorporaram os conceitos de indiscernibilidade e dominância para a indução de regras decisórias em contextos multicritério (BŁASZCZYŃSKI, SLOWINSKI e SZELAG, 2011). Posteriormente, métodos como o VC-DRSA foram introduzidos para flexibilizar a definição da aproximação inferior, permitindo a inclusão controlada de contraexemplos e aumentando a adaptabilidade aos dados ruidosos e às incertezas inerentes à modelagem decisória (BŁASZCZYŃSKI et al., 2006). Essa evolução teórica foi complementada pela integração dos *Rough Sets* com outras técnicas de inteligência artificial, como lógica *fuzzy*, redes neurais e algoritmos evolutivos que juntos ampliaram o leque de aplicações para problemas de classificação, ranking e avaliação de riscos, evidenciando a

robustez e a versatilidade dessas abordagens em cenários complexos (SLOWINSKI e SZELAG, 2011).

Dentro desse cenário, o algoritmo PREPPY (*Permutation-based Recall and Precision Pattern Yielding*) se destaca ao incorporar os sólidos fundamentos da DRSA e, ao mesmo tempo, adotar uma estratégia inovadora baseada na ordenação permutacional dos atributos. Diferentemente de métodos tradicionais, como o DomLEM, o PREPPY explora o ganho de informação para hierarquizar atributos e ajusta dinamicamente os limiares de decisão, possibilitando a indução de regras concisas, interpretáveis e adaptadas tanto para critérios quanto atributos categóricos. Essa abordagem não só melhora o *recall* e *precision* ao enfrentar a heterogeneidade dos dados, mas também enfatiza a análise incremental e o controle do *trade-off* entre abrangência e especificidade das regras, estabelecendo assim uma metodologia única e promissora para a classificação em contextos multicritério.

3. PREPPY

No cenário da análise de decisão multicritério, a constante busca por algoritmos capazes de lidar com a complexidade e heterogeneidade dos dados impulsiona o desenvolvimento de novas abordagens. Nesse contexto, o algoritmo PREPPY (*Permutation-based Recall and Precision Pattern Yielding*) se destaca ao incorporar a robusta *Dominance-Based Rough Set Approach* (DRSA) na exploração da ordenação de atributos para indução de regras de maneira incremental.

O algoritmo é compatível tanto com critérios, considerados atributos numéricos ordinais, quanto com atributos categóricos. Por sua vez, atributos categóricos são definidos como atributos em uma escala qualitativa, distinguindo-se entre ordinais e nominais. Enquanto que os atributos categóricos ordinais induzem relações de preferência, os atributos nominais – isto é, os categóricos sem ordenação – não são capazes de estabelecer relações de preferência. Essa adaptação permitiu não só tratar os critérios com preferências de ganho e custo de maneira tradicional, mas também incorporar atributos categóricos de dois tipos:

- Atributos categóricos ordinais: Possuem relações de ordem entre as categorias (por exemplo, “>” ou “<”, além de “=” e “≠”). Com esses atributos, é possível estabelecer relações de dominância, embora a escala possa apresentar empates.
- Atributos categóricos não ordinais: Apresentam apenas relações de igualdade ou desigualdade (“=” e “≠”), o que impede a definição de que uma categoria seja melhor ou pior que outra, pois elas apenas empatam.

Portanto, o PREPPY foi desenvolvido primordialmente para escalas ordinais, onde a existência de pelo menos uma relação de ordem permite a construção de relações de dominância com base nas comparações entre categorias. Embora o algoritmo aceite algum atributo categórico não ordinal, este não é o foco, pois, para tais atributos, não é possível formar relações de dominância que indiquem uma hierarquia.

A estrutura do PREPPY organiza os dados em uma matriz de decisão, acompanhada de um vetor de preferências para os critérios e de matrizes de relações entre categorias para os atributos categóricos. Assim como na DRSA, o algoritmo estabelece relações de dominância entre os objetos, forma uniões de classes e calcula as aproximações inferior e superior, definindo também a região de fronteira (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011; GRECO et al., 2000, 2001). Cada objeto é representado por um vetor de

atributos e associado a uma classe de decisão. Os critérios são classificados como de ganho (onde valores mais altos são desejáveis) ou de custo (priorizando valores menores), enquanto os atributos categóricos, quando ordinais, são utilizados para formar relações de dominância com base nas relações preestabelecidas entre categorias.

Além disso, o algoritmo utiliza o cálculo do ganho de informação para identificar os atributos mais discriminantes e estabelecer a ordem de permutação dos critérios durante o ajuste dos parâmetros. O conjunto de treino depende do julgamento do decisor para definir as classes, seja com base em decisões passadas ou por meio de uma análise holística do desempenho dos objetos nos atributos.

Com isso, o PREPPY emprega a DRSA para induzir regras de classificação que abrangem tanto os critérios quanto os atributos categóricos. Essa abordagem permite extrair conhecimento dos objetos que compõem a aproximação inferior das classes, proporcionando uma robusta capacidade preditiva para contextos multicritério.

Diferenciando-se de abordagens como o DomLEM – que prioriza a indução de um conjunto mínimo de regras consistentes com o princípio da dominância em critérios (BLASZCZYNSKI, SLOWINSKI e SZELAG, 2011) – o PREPPY aposta na exploração das ordenações dos atributos para induzir regras mais concisas e interpretáveis. Essa ênfase no desempenho classificatório na generalização, método único de indução de regras e na capacidade de tratar dados heterogêneos posiciona o PREPPY de forma singular no campo da análise de decisão baseada em DRSA.

Contudo, o PREPPY apresenta limitações importantes em termos de complexidade computacional, especialmente devido à análise exaustiva das permutações dos critérios para ajuste dos parâmetros, o que pode comprometer a escalabilidade em problemas com muitos atributos. Além disso, seu desempenho depende da assertividade das matrizes de relações para atributos categóricos; se essas matrizes não refletirem adequadamente o conhecimento do domínio, o desempenho do modelo pode ser prejudicado. Há ainda certa sensibilidade a dados ruidosos e *outliers*, que podem levar ao *overfitting*, e proliferação de regras geradas em contextos de alta heterogeneidade, dificultando a interpretação e aplicação prática do modelo. Para mitigar esses desafios e equilibrar métricas como *precision* e *recall*, pretende-se incluir uma análise de sensibilidade que permita relaxar dinamicamente as regras geradas – ajustando limiares de confiança e cobertura – de modo a otimizar o *trade-off* entre *precision* e abrangência das predições.

A seguir, detalham-se os conceitos e a aplicação prática do algoritmo, desde a definição do universo de objetos e critérios até a formulação das regras de decisão que permitem a classificação de novos objetos, demonstrando como o PREPPY enfrenta os desafios da análise de decisão multicritério com uma abordagem que integra, de forma nativa, a utilização de atributos categóricos ordinais para a formação de relações de dominância.

3.1. Descrição Conceitual das Etapas

Seja U o conjunto universal de objetos x (alternativas) e $A = \{a_1, a_2, \dots, a_n\}$ o conjunto de atributos utilizados para a avaliação. Cada objeto $x_i \in U$ é descrito por um vetor de características e uma classe de decisão, isto é:

$$[42] \quad x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n}, d_i),$$

Onde d_i denota a classe de x_i . Para cada critério a_j , existe uma preferência associada que pode ser de ganho – onde valores maiores são desejáveis, de custo – onde valores menores são desejáveis – e categórico ordinal ou não – em que os valores pertencem a um conjunto de categorias cuja relação é definida por operadores relacionais (GRECO et al., 2000).

O sistema de informação é construído a partir de uma matriz de decisão \mathcal{M} em que cada linha corresponde a um objeto e cada coluna a um critério, além da coluna de decisão. Formalmente, podemos escrever:

$$[43] \quad \mathcal{M} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} & d_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{m,1} & \cdots & x_{m,n} & d_m \end{bmatrix}$$

As preferências dos critérios são registradas em um vetor $P = (p_1, p_2, \dots, p_n)$, onde $p_i \in \{\text{ganho}, \text{custo}, \text{categórico}\}$ (GRECO et al., 2000).

Para atributos categóricos, é definida uma matriz de relações pares $R^{(a)}$ para cada atributo a do conjunto $A_{\text{categórico}}$. Essa matriz relaciona cada par de categorias (v, w) segundo um operador relacional $\rho \in \{>, <, =, \neq\}$, isto é:

$$[44] \quad R^{(a)}(v, w) = \rho$$

A ideia central do PREPPY é fundamentada na *Dominance-Based Rough Set Approach* (DRSA) do Greco et al. (2000), onde a relação de dominância é definida da seguinte

forma: Para dois objetos $x, y \in U$, diz-se que x domina y (simbolizado por $x \succ y$) se, para todo critério $a_i \in A$, vale que:

$$[45] \quad \begin{cases} x_i \geq y_i, & \text{se } p_i = \text{ganho} \\ x_i \leq y_i, & \text{se } p_i = \text{custo} \end{cases}$$

e exista ao menos um índice j tal que a desigualdade seja estrita, isto é:

$$[46] \quad \exists j \in \{1, \dots, n\}: \begin{cases} x_j > y_j, & \text{se } p_j = \text{ganho}, \\ x_j < y_j, & \text{se } p_j = \text{custo} \end{cases}$$

Para atributos categóricos, a comparação entre os valores v e w é realizada utilizando a matriz $R^{(a)}$.

Após o estabelecimento das relações de dominância, são construídas duas uniões de classes. Seja $\{C_1, C_2, \dots, C_m\}$ o conjunto das classes de decisão ordenadas de forma ascendente – da mais baixa à mais alta. A união de classes ascendentes a qual agrupa os objetos que pertencem a classes inferiores ou iguais à uma classe C_k :

$$[47] \quad C_k^{\geq} = \bigcup_{j=1}^k C_j,$$

E também segundo Greco et al. (2000), a união de classes descendentes agrupa os objetos que pertencem à C_k ou classes superiores:

$$[48] \quad C_k^{\leq} = \bigcup_{j=k}^m C_j$$

Dada uma das uniões de classes, o algoritmo define duas aproximações, superior e inferior. Segundo Greco et al. (2000), a aproximação inferior \underline{C} , com o conjunto de objetos que, com base na relação de dominância, pertencem inequivocamente à classe C :

$$[49] \quad \underline{C}_k^{\geq} = \{x \in U: D_P^+(x) \subseteq C_k^{\geq}\}$$

$$[50] \quad \underline{C}_k^{\leq} = \{x \in U: D_P^-(x) \subseteq C_k^{\leq}\}$$

E a aproximação superior \overline{C}^+ , com o conjunto de objetos que possivelmente pertencem à classe C , isto é, existem evidências na relação de dominância que os desassocia de C :

$$[51] \quad \overline{C}_k^{\geq} = \bigcup_{x \in C_k^{\geq}} D_P^+(x)$$

$$[52] \quad \bar{C}_k^{\leq} = \bigcup_{x \in C_k^{\leq}} D_P^-(x)$$

Consequentemente, a região de fronteira é definida por:

$$[53] \quad B(C) = \bar{C} \setminus \underline{C}$$

São calculadas as aproximações inferiores e superiores tanto para a união de classes ascendente como descendente, ou seja, ao total quatro aproximações.

Antes da indução de regras, o sistema determina quais atributos (tanto os critérios numéricos quanto os categóricos) possuem maior relevância para a classificação, utilizando o ganho de informação. O ganho de informação é uma métrica da teoria da informação que quantifica a redução da incerteza sobre a variável alvo – neste caso a classe de decisão – ao conhecermos o valor de um atributo:

$$[54] \quad H(d) = - \sum_{k=1}^m p(d_k) \log_2(p(d_k))$$

Onde $H(d)$ representa a entropia associada ao atributo de decisão d , m é o número de classes de decisão, $p(d_k)$ é a probabilidade decorrência da classe d_k e \log_2 mede a informação em bits. Também é necessário calcular a entropia condicional para cada um dos atributos:

$$[55] \quad H(d|a_i) = - \sum_{v \in V_i} p(a_i = v) \sum_{k=1}^m p(d_k|a_i = v) \log_2(p(d_k|a_i = v))$$

Onde $H(d|a_i)$ representa a entropia condicional da decisão d dado o atributo a_i , V_i é o conjunto de valores possíveis para o atributo a_i , $p(a_i = v)$ é a probabilidade de que o atributo a_i assuma o valor v e $p(d_k|a_i = v)$ é a probabilidade condicional de que a decisão seja d_k dado que a_i tenha o valor de v . Para finalizar a ordenação das permutações, é calculada o ganho de informação como:

$$[56] \quad \text{Ganho de Informação} = H(d) - H(d|a_i)$$

Em que atributos com maior ganho de informação são considerados mais discriminatórios. Após o cálculo, atributos são classificados em ordem decrescente de ganho de informação, garantindo que os mais relevantes sejam priorizados na etapa de indução de regras e considerando primeiro atributos com maior poder preditivo.

Com os atributos ordenados conforme sua relevância, o sistema inicia a indução de regras utilizando os objetos na aproximação inferior da união de classes descendentes. Para cada critério a_i , calcula-se o número de casas decimais s_i dos valores dos objetos da aproximação inferior. A partir disso, define-se o passo de ajuste:

$$[57] \quad \Delta_i = 10^{-s_i}$$

O algoritmo explora as permutações dos critérios para garantir a indução de regras que classifiquem adequadamente todos os objetos de uma dada classe da união de classes. Seja Π o conjunto de todas as permutações dos critérios $\{1, 2, \dots, n\}$. Para cada permutação $\pi \in \Pi$, o algoritmo ordena limiares de critérios:

$$[58] \quad \theta_{\pi(1)}, \theta_{\pi(2)}, \dots, \theta_{\pi(n)}$$

Definidos como:

$$[59] \quad \theta_{\pi i}^{(0)} = \begin{cases} \min\{x_{\pi(i)} | x \in C^-\}, & \text{se } p_{\pi(i)} = \text{ganho}, \\ \max\{x_{\pi(i)} | x \in C^-\}, & \text{se } p_{\pi(i)} = \text{custo} \end{cases}$$

E os ajusta conforme:

$$[60] \quad \theta_{\pi(i)}^{(t+1)} = \theta_{\pi(i)}^{(t)} + \delta_{\pi(i)},$$

Onde $\delta_{\pi(i)} = -\Delta_{\pi(i)}$ para ganho e $\delta_{\pi(i)} = \Delta_{\pi(i)}$ para custo.

À cada iteração, o conjunto de objetos que satisfaz a regra é avaliado por *precision* e *recall*:

$$[61] \quad \textit{precision} = \frac{n^\circ \text{ de objetos corretamente classificados}}{n^\circ \text{ total de objetos cobertos pela regra}}$$

$$[62] \quad \textit{recall} = \frac{n^\circ \text{ de objetos corretamente classificados}}{n^\circ \text{ total de objetos pertencentes à classe}}$$

Uma regra típica R tem a forma:

$$[63] \quad R: \bigwedge_{i=1}^n (x_i \sim_i \theta_i) \rightarrow d = k,$$

Onde θ_i é um limiar para o critério a_i , \sim_i é o operador de comparação definido como \geq se $p_i = \text{ganho}$ ou \leq se $p_i = \text{custo}$. Para atributos categóricos, a regra pode assumir a forma:

$$[64] \quad R': \bigwedge_{a \in A_{\text{categ\u00f3rico}}} (x_a \in \{v_{a,1}, v_{a,2}, \dots\}) \rightarrow d = k,$$

Em que $A_{\text{categ\u00f3rico}} \subset A$ s\u00e3o os atributos categ\u00f3ricos. A regra r \u00e9 ajustada iterativamente at\u00e9 que se satisfa\u00e7am uma das duas condi\u00e7\u00f5es:

$$[65] \quad \textit{precision}(r) = 1 \ \& \ \textit{recall}(r) = 1$$

$$[66] \quad \Delta_{\textit{precision}}(r) \leq 0 \ \text{ou} \ \Delta_{\textit{recall}}(r) \leq 0$$

Em termos matem\u00e1ticos, o ajuste \u00e9 realizado enquanto:

$$[67] \quad \begin{cases} \textit{precision}(r) < 1 \ \text{ou} \ \textit{recall}(r) < 1 \\ \text{e} \ \Delta_{\textit{precision}}(r) > 0, \quad \text{com} \ \Delta_{\textit{recall}}(r) \end{cases}$$

Caso o *recall* n\u00e3o atinja o valor de 1, o algoritmo continua explorando as permuta\u00e7\u00f5es dos crit\u00e9rios at\u00e9 que se alcance a condi\u00e7\u00e3o desejada ou que se esgotem as possibilidades.

Ap\u00f3s a fase de permuta\u00e7\u00f5es, \u00e9 necess\u00e1rio filtrar as regras para reter apenas as \u00f3timas. Seja \mathcal{R} o conjunto de todas as regras geradas. Em seguida, procede-se da seguinte forma:

1. Se existir alguma regra $r^* \in \mathcal{R}$ com $\textit{recall}(r^*) = 1$, essa regra \u00e9 escolhida como \u00fanica representante;
2. Caso nenhuma regra atinja $\textit{recall} = 1$ ap\u00f3s todas as permuta\u00e7\u00f5es, as regras s\u00e3o inicialmente ordenadas de acordo com a *precision*. Ou seja, para duas regras r_i e r_j , a ordena\u00e7\u00e3o segue:

$$[68] \quad \textit{precision}(r_i) \geq \textit{precision}(r_j);$$

3. Regras duplicadas s\u00e3o removidas;
4. Para eliminar regras dominadas, compara-se cada par de regras r_i e r_j em todos os atributos dando prefer\u00eancia por regras mais abrangentes. \u00c9 considerado que $r_i D r_j$ se r_i for mais abrangente que r_j em ao menos um atributo e t\u00e3o abrangente quanto em todos os demais.

Ap\u00f3s percorrer todas as combina\u00e7\u00f5es poss\u00edveis, o conjunto final \mathcal{R}^* conter\u00e1 apenas as regras n\u00e3o dominadas, que otimizam de forma balanceada as m\u00e9tricas de *precision* e *recall*.

Para classificar um novo objeto $x = (x_1, x_2, \dots, x_n)$ o sistema verifica quais regras R e R' s\u00e3o satisfeitas:

[69] Se $x_i \sim_i \theta_i$ para todos i (e $x_a \in \{v_{a,1}, v_{a,2}, \dots\}$ para atributos categóricos)

Então o objeto é classificado na classe $d = k$ indicada pela regra. Quando múltiplas regras se aplicam, são priorizadas regras com maior *precision* da classe de maior ordem. Após a classificação, são calculadas as métricas para avaliar seu desempenho.

3.2. Métricas de Classificação

O PREPPY adicionalmente produz duas matrizes de confusão complementares – *Predicted vs Original* (PvO), que confronta diretamente as previsões do modelo com as classes verdadeiras, e *Predicted vs Approximation* (PvA), que verifica a concordância entre cada predição e a região de aproximação definida pela DRSA – permitindo diagnosticar de forma granular erros como falsos positivos e falsos negativos (OPITZ, 2024).

Para a avaliação formal do desempenho do PREPPY é definindo, para cada classe de decisão $C_k \subseteq \{C_1, \dots, C_m\}$, os quatro elementos da matriz de confusão da predição em relação às classes original sobre o universo U :

$$[70] \quad TP_k = |\{x \in U | f(x) = C_k \wedge d(x) = C_k\}|$$

$$[71] \quad FP_k = |\{x \in U | f(x) = C_k \wedge d(x) \neq C_k\}|$$

$$[72] \quad FN_k = |\{x \in U | f(x) \neq C_k \wedge d(x) = C_k\}|$$

$$[72] \quad TN_k = |U| - (TP_k + FP_k + FN_k)$$

Respectivamente *True Positive* – objetos corretamente classificados na classe de interesse, *False Positive* – objetos incorretamente classificados na classe de interesse, *False Negative* – objetos da classe de interesse incorretamente classificados em outra – e *True Negative* – objetos corretamente classificados como não pertencentes à classe de interesse (OPITZ, 2024).

A partir dessas contagens, são definidas as métricas principais de avaliação do PREPPY. Cada uma dessas métricas é calculada em três variantes – *Macro average* (média simples entre classes), *Micro average* (cálculo global considerando todos os exemplos) e *Weighted average* (média ponderada pelo número de exemplos em cada classe) – permitindo avaliar tanto o desempenho geral quanto o comportamento específico em cenários de classes desbalanceadas (OPITZ, 2024). Será definido à princípio a acurácia, que mensura a proporção de previsões corretas sobre o total de casos, fornecendo uma visão geral do desempenho do modelo:

$$[74] \quad Accuracy = \frac{\sum_{k=1}^m TP_k}{|U|}$$

Então a *precision*, que indica a proporção de verdadeiros positivos entre todas as predições positivas feitas pelo modelo:

$$[75] \quad Precision_{macro} = \frac{1}{m} \sum_{k=1}^m \frac{TP_k}{TP_k + FP_k}$$

$$[76] \quad Precision_{micro} = \frac{\sum_k TP_k}{\sum_k (TP_k + FP_k)}$$

$$[77] \quad Precision_{weighted} = \sum_{k=1}^m \frac{|C_k|}{|U|} \frac{TP_k}{TP_k + FP_k}$$

O *recall*, que reflete a capacidade do modelo de identificar corretamente todos os exemplos pertencentes a uma classe:

$$[78] \quad Recall_{macro} = \frac{1}{m} \sum_{k=1}^m \frac{TP_k}{TP_k + FN_k}$$

$$[79] \quad Recall_{micro} = \frac{\sum_k TP_k}{\sum_k (TP_k + FN_k)}$$

$$[80] \quad Recall_{weighted} = \sum_{k=1}^m \frac{|C_k|}{|U|} \frac{TP_k}{TP_k + FN_k}$$

o *F1-Score*, que combina *precision* e *recall* em um único valor harmônico:

$$[81] \quad F1_{macro} = \frac{1}{m} \sum_{k=1}^m 2 \frac{Precision_k \cdot Recall_k}{Precision_k + Recall_k}$$

$$[82] \quad F1_{micro} = 2 \frac{Precision_k \cdot Recall_k}{Precision_k + Recall_k}$$

$$[83] \quad F1_{weighted} = \sum_{k=1}^m \frac{|C_k|}{|U|} 2 \frac{Precision_k \cdot Recall_k}{Precision_k + Recall_k}$$

Como métricas complementares, são calculados o *Matthews Correlation Coefficient* (MCC), que oferece uma medida balanceada de qualidade de classificação mesmo em situações de forte desequilíbrio de classes:

$$[85] \quad MCC_k = \frac{TP_k \cdot TN_k - FP_k \cdot FN_k}{\sqrt{(TP_k + FP_k)(TP_k + FN_k)(TN_k + FP_k)(TN_k + FN_k)}}$$

E o *Cohen's Kappa*, que quantifica o nível de concordância entre previsões e classes reais corrigindo-se pela chance:

$$[86] \quad P_o = \frac{TP_k + TN_k}{|U|}$$

$$[87] \quad \frac{(TP_k + FP_k)(TP_k + FN_k) + (TN_k + FP_k)(TN_k + FN_k)}{|U|^2}$$

$$[88] \quad \kappa_k = \frac{P_o - P_e}{1 - P_e}.$$

O uso combinado dessas métricas garante uma avaliação abrangente da performance do PREPPY. Enquanto as métricas principais fornecem uma visão rápida e comparável de acurácia e cobertura das classes, as métricas complementares aprofundam a análise ao revelar vieses, robustez frente a desequilíbrios e a confiabilidade geral do modelo. Essa abordagem multifacetada de avaliação é essencial para validar a eficácia preditiva e a aplicabilidade prática do PREPPY em contextos multicritério complexos (OPITZ, 2024).

3.3. Implementação do PREPPY

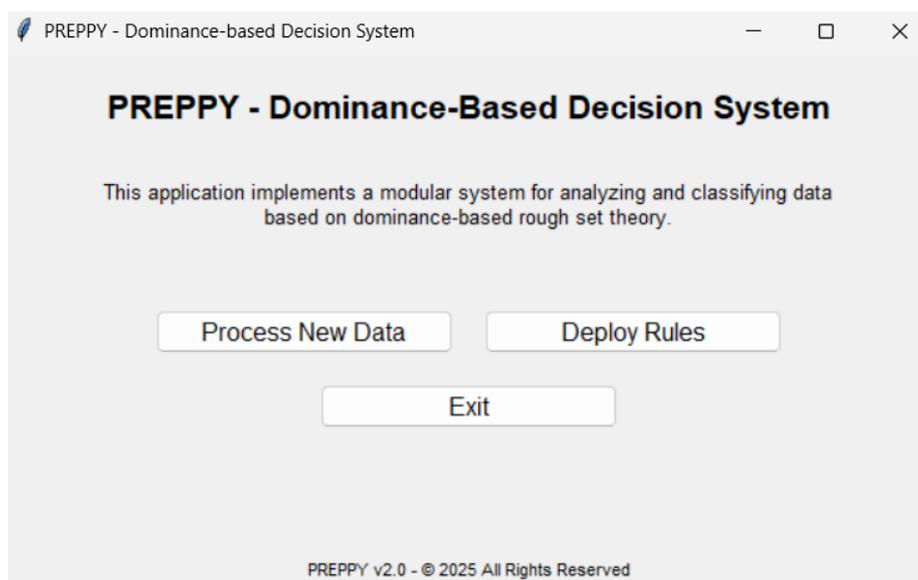
A implementação do PREPPY em código é estruturada de forma modular, iniciando sua execução por meio de um ponto central que configura o ambiente, processa os argumentos da linha de comando e aciona a execução do sistema, permitindo sua utilização tanto via interface gráfica quanto por linha de comando. O código está organizado em dois grandes conjuntos: o *back-end*, que concentra toda a lógica de negócio, processamento e indução de regras com base na *Dominance-Based Rough Set Approach* (DRSA), e o *front-end*, responsável por fornecer uma interface amigável para interação do usuário, facilitando desde a entrada e saída dos dados até a visualização dos resultados.

No lado do *back-end*, encontram-se os módulos que lidam com a importação, limpeza e exportação dos dados, a construção do sistema de informação (organizando objetos e atributos, seja critérios ou categóricos), o cálculo de aproximações inferiores e superiores e a indução de regras através de técnicas que exploram permutações dos critérios e ajustam limiares com base na sensibilidade de cada atributo. Além disso, há funções dedicadas ao tratamento de atributos categóricos, geração de matrizes de relações e à avaliação da performance do modelo

por meio de métricas de classificação e matrizes de confusão. Por sua vez, o *front-end* agrupa funcionalidades de interface que englobam o menu principal, telas para o carregamento e exportação dos dados, interfaces específicas para o processamento e *deploy* dos dados e ferramentas de visualização, que auxiliam o usuário na análise e interpretação dos resultados obtidos.

A implementação do PREPPY destaca-se pelo compromisso com a alta performance e a manutenção facilitada. O sistema utiliza operações vetorizadas para acelerar comparações e cálculos, aliadas a um gerenciamento de memória aprimorado em operações intensivas. Essa estratégia, combinada à clara separação entre *back-end* e *front-end*, possibilita que alterações e melhorias sejam testadas, mantidas e expandidas de forma prática, garantindo assim um sistema robusto, flexível e escalável para futuras evoluções. Todo o desenvolvimento do PREPPY é gerenciado em um repositório, no momento privado, no GitHub, o que assegura uma gestão de versionamento rigorosa, backups regulares e um histórico detalhado de alterações, facilitando possíveis colaborações entre desenvolvedores e a implementação contínua de aprimoramentos.

Figura 1 – Interface do PREPPY



Fonte: O autor (2025)

Além disso, o PREPPY foi concebido para ser uma ferramenta acessível a decisores de diversas áreas – desde gestores e analistas até profissionais que necessitam de suporte prático para a tomada de decisões multicritério – dispensando conhecimentos técnicos aprofundados em computação. Sua interface intuitiva e estrutura modular abstraem a complexidade dos algoritmos subjacentes, permitindo que o usuário se concentre na

interpretação dos resultados e na identificação de oportunidades de melhoria, sem a preocupação com os detalhes da implementação. Dessa forma, o sistema se mostra adequado tanto para especialistas que desejam aprofundar sua análise quanto para usuários que buscam uma solução robusta e de fácil uso para embasar suas decisões.

3.4. Conjunto de Exemplo

Para melhor compreensão do PREPPY, será utilizado como exemplo o conjunto de dados *evaluation in a high school* proposto por Grabisch em 1994 e alterado por Greco, Matarazzo e Slowinski em 1999 ao introduzir mais nove instâncias e modificar alguns valores dos atributos, sobretudo o de literatura e de decisão, os quais passaram a representarem uma escala de três valores ao invés do original de duas. Além disso, o atributo da classe de decisão foi discretizado para compatibilização com o PREPPY, isto é, seus valores qualitativos foram transformados em quantitativos, mas mantendo a mesma escala. O exemplo apresentado descreve uma situação em que o diretor de uma escola deseja classificar alunos em três turmas – *Good*, *Medium* e *Bad* – com base em seu desempenho em três disciplinas (Matemática, Física e Literatura).

Tabela 1 – Exemplos do conjunto *evaluation in a high school*

<i>Students</i>	<i>Math</i>	<i>Physics</i>	<i>Literature</i>	<i>Global Evaluation</i>
1	<i>Bad</i>	<i>Good</i>	<i>Bad</i>	1
2	<i>Good</i>	<i>Medium</i>	<i>Bad</i>	2
3	<i>Medium</i>	<i>Good</i>	<i>Good</i>	3
4	<i>Medium</i>	<i>Medium</i>	<i>Good</i>	2
5	<i>Good</i>	<i>Good</i>	<i>Bad</i>	2
6	<i>Medium</i>	<i>Medium</i>	<i>Bad</i>	1
7	<i>Good</i>	<i>Bad</i>	<i>Medium</i>	1
8	<i>Good</i>	<i>Medium</i>	<i>Medium</i>	3
9	<i>Good</i>	<i>Medium</i>	<i>Good</i>	3
10	<i>Medium</i>	<i>Good</i>	<i>Bad</i>	2
11	<i>Bad</i>	<i>Bad</i>	<i>Medium</i>	1
12	<i>Medium</i>	<i>Good</i>	<i>Medium</i>	3
13	<i>Good</i>	<i>Bad</i>	<i>Good</i>	1
14	<i>Medium</i>	<i>Medium</i>	<i>Medium</i>	2
15	<i>Bad</i>	<i>Good</i>	<i>Good</i>	1

Fonte: O autor (2025)

O qual foi dividido de maneira à princípio aleatória em *training* contendo as sete primeiras alternativas e *deployment* contendo as restantes. Foi buscado garantir uma certa distribuição mínima dos valores em ambos os conjuntos de *training* e *deployment* com fins de induzir regras sólidas a serem aplicadas em um conjunto também condizente. Após a construção do sistema de informações com a matriz de desempenho, o vetor de preferências e as matrizes de relações entre categorias, exemplificada na Tabela 2, que por padrão foram estabelecidas como *Good*>*Medium*>*Bad* para todos os atributos categóricos por se tratar de uma escala estabelecida. E a construção das uniões ascendentes, isto é, da classe atual para cima, e descendentes, da classe atual para baixo na Tabela 3.

Tabela 2 – Matriz de relações entre categorias

	<i>Bad</i>	<i>Medium</i>	<i>Good</i>
<i>Bad</i>	=	<	<
<i>Medium</i>	>	=	<
<i>Good</i>	>	>	=

Fonte: O autor (2025)

Tabela 3 – Uniões ascendentes e descendentes

	Upward Unions	Downward Unions
<i>Class 1</i>	[1, 2, 3, 4, 5, 6, 7]	[3, 5, 7]
<i>Class 2</i>	[1, 2, 4, 6]	[2, 3, 5, 6, 7]
<i>Class 3</i>	[1, 4]	[1, 2, 3, 4, 5, 6, 7]

Fonte: O autor (2025)

O algoritmo então parte para as definições das relações de dominância, isto é, as relações de objetos que são dominados por outros e objetos que dominam outros do conjunto, estabelecendo a Tabela 4 e o cálculo das aproximações inferior, superior e da região de fronteira na Tabela 5.

Tabela 4 –Relações de dominância

<i>Object</i>	1	2	3	4	5	6	7
<i>Dominates</i>	[3, 5, 6]	[]	[]	[2, 3, 6]	[3]	[3]	[3]
<i>Dominated by</i>	[]	[4]	[1, 4, 5, 6, 7]	[]	[1]	[1, 4]	[]

Fonte: O autor (2025)

Tabela 5 – Aproximações superior, região de fronteira e inferior da união de classes ascendente

<i>Class</i>	1	2	3
<i>Upper</i>	[1, 2, 3, 4, 5, 6, 7]	[1, 2, 4, 6]	[1, 4]
<i>Boundaries</i>	[]	[]	[]
<i>Lower</i>	[1, 2, 3, 4, 5, 6, 7]	[1, 2, 4, 6]	[1, 4]

Fonte: O autor (2025)

As quais ambas são iguais pois não existem inconsistências ou ambiguidades nos dados tratando-se das relações de dominância estabelecidas. Nos sistemas de decisão baseados em dominância, como o PREPPY, as aproximações refletem o grau de certeza com que objetos pertencem a determinadas classes. Quando estas aproximações coincidem, significa que o conjunto de dados apresenta perfeita consistência relacional – cada objeto é categorizado precisamente de acordo com suas relações de dominância, sem objetos conflitantes que possam características similares mas classificações distintas. Este fenômeno indica alta qualidade dos atributos de condição selecionados para representar o problema de decisão, pois eles capturam perfeitamente as características que determinam a classificação, resultando em fronteiras de decisão nítidas e sem regiões limítrofes ambíguas entre as classes. Fora calculado o ganho de informação para cada um dos atributos, estabelecendo a ordem de priorização dos atributos na indução de regras e permutação como na Tabela 6.

Tabela 6 – Resultados dos cálculos de ganho de informação

	<i>Math</i>	<i>Physics</i>	<i>Literature</i>
<i>Information Gain</i>	0,8774	0,5917	0,4838
<i>Rank</i>	1°	2°	3°

Fonte: O autor (2025)

No caso deste conjunto de dados contendo apenas atributos categóricos ordinais, as regras foram induzidas de maneira descritiva, capturando todos os valores presentes em cada uma das classes de maneira intercalada. Como em nenhum momento foi estabelecido recall e *precision* de 100%, as descrições dos valores ocorreram para todos os atributos na terceira classe. Para o caso da segunda classe, fora calculadas diversas permutações e definida como melhor regra o conjunto evidenciada na Tabela 7.

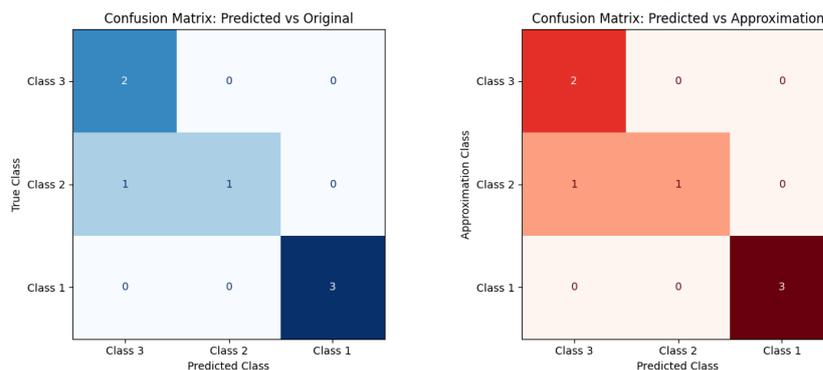
Tabela 7 – Regras induzidas

<i>Class</i>	<i>Math</i>	<i>Physics</i>	<i>Literature</i>	<i>Recall</i>	<i>Precision</i>
3	<i>Medium or Good</i>	<i>Medium or Good</i>	<i>Medium or Good</i>	100,00%	66,67%
2	<i>Medium or Good</i>	<i>Medium or Good</i>	<i>Any</i>	50,00%	100,00%
1	<i>Any</i>	<i>Any</i>	<i>Any</i>	100,00%	100,00%

Fonte: O autor (2025)

No sistema PREPPY, este o fenômeno de sobreposição de antecedentes explica a discrepância entre o *recall* teórico (100%) e o *recall* efetivo (50%) para a classe 2. Enquanto a indução de regras sugere uma cobertura perfeita quando avaliada isoladamente, na aplicação sequencial das regras no conjunto completo, observa-se uma competição entre regras com condições similares. Como o algoritmo prioriza a primeira regra aplicável encontrada, seguindo a ordem de processamento, ocorre a captura indevida de instâncias que satisfazem ambas as regras pela classe de maior prioridade. Essa característica evidencia que a especificidade e a ordem das regras determinam a eficácia do modelo, impactando métricas como o recall nas classes seguintes. Fora gerado matrizes de confusão e métricas calculadas para o desempenho geral e cada uma das classes são determinadas, como visto nas Tabelas 8 e 9.

Figura 2 – Matrizes de confusão previsto com original e previsto com aproximação do treinamento



Fonte: O autor (2025)

Tabela 8 – Métricas gerais do treinamento

<i>Metric</i>	<i>Value</i>
<i>Accuracy</i>	75,0%
<i>Recall (Macro)</i>	77,8%
<i>Recall (Micro)</i>	75,0%
<i>Recall (Weighted)</i>	75,0%
<i>Precision (Macro)</i>	77,8%
<i>Precision (Micro)</i>	75,0%

<i>Precision (Weighted)</i>	79,2%
<i>F1-Score (Macro)</i>	75,6%
<i>F1-Score (Micro)</i>	75,0%
<i>F1-Score (Weighted)</i>	75,0%
<i>Matthews Correlation Coefficient (MCC)</i>	64,3%
<i>Cohen's Kappa</i>	62,8%

Fonte: O autor (2025)

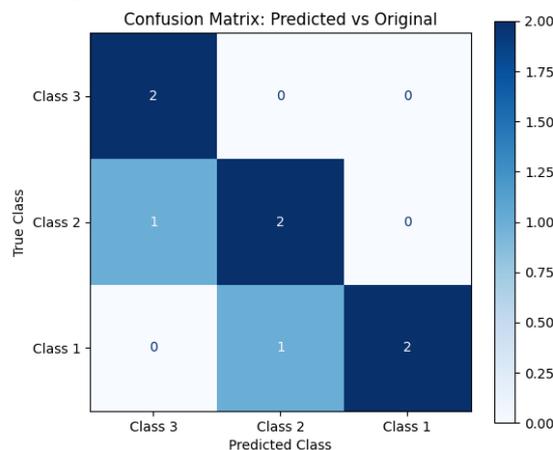
Tabela 9 – Métricas por classe do treinamento

	<i>Recall</i>	<i>Precision</i>	<i>F1-Score</i>
<i>Class 3</i>	100,00%	66,67%	80,00%
<i>Class 2</i>	50,00%	100,00%	66,67%
<i>Class 1</i>	100,00%	100,00%	100,00%

Fonte: O autor (2025)

De modo geral, a acurácia de 75% indica que, em média, três a cada quatro predições estão corretas. Observa-se um excelente desempenho na classe 1 (100% de *precision* e *recall*), mas há certa dificuldade na classe 2, onde a *precision* é alta (100%), porém o *recall* é baixo (50%), sugerindo que a metade dos exemplos dessa classe não foi corretamente identificada. Já a classe 3 atinge 100% de *recall*, mas uma *precision* de 66,67%, o que implica em algumas predições incorretas como sendo classe 3. As métricas globais (F1 macro de 75,6%, MCC de 64,3% e Kappa de 62,8%) apontam para um modelo razoavelmente equilibrado, porém com margem para melhorias, especialmente na detecção de exemplos da classe 2. Segundo assim, essas regras induzidas através do conjunto de *training* foram utilizadas para classificar o conjunto de *deployment*, gerando a seguinte matriz de confusão:

Figura 3 – Matriz de confusão do *deployment*



Fonte: O autor (2025)

Por fim, foram calculadas as métricas que representam a capacidade de generalização das regras induzidas pelo PREPPY vistas nas Tabelas 10 e 11.

Tabela 10 – Métricas gerais do *deployment*

<i>Metric</i>	<i>Value</i>
<i>Accuracy</i>	85,71%
<i>Recall (Macro)</i>	83,33%
<i>Recall (Micro)</i>	85,71%
<i>Recall (Weighted)</i>	85,71%
<i>Precision (Macro)</i>	88,89%
<i>Precision (Micro)</i>	85,71%
<i>Precision (Weighted)</i>	90,48%
<i>F1-Score (Macro)</i>	82,22%
<i>F1-Score (Micro)</i>	85,71%
<i>F1-Score (Weighted)</i>	84,76%
<i>Matthews Correlation Coefficient (MCC)</i>	80,69%
<i>Cohen's Kappa</i>	78,13%

Fonte: O autor (2025)

Tabela 11 – Métricas por classe do *deployment*

	<i>Recall</i>	<i>Precision</i>	<i>F1-Score</i>
<i>Class 3</i>	100,00%	66,67%	80,00%
<i>Class 2</i>	66,67%	66,67%	66,67%
<i>Class 1</i>	66,67%	100,00%	80,00%

Fonte: O autor (2025)

No geral, a acurácia de 85,71% no conjunto de *deployment* indica que o modelo manteve um bom desempenho ao generalizar para novos dados. Além disso, métricas como *Precision (Micro)* e *Recall (Micro)* igualmente em 85,71% sugerem uma consistência entre a proporção de exemplos preditos corretamente e a capacidade de recuperar instâncias positivas em cada classe. Os valores de MCC (80,69%) e Cohen's Kappa (78,13%) também reforçam uma boa concordância entre as predições e os valores reais, indo além de um acerto ao acaso. Esses números globais sinalizam que o modelo tem bom potencial para aplicação prática, embora ainda haja variações de desempenho entre as classes.

Em nível de classe, observa-se que a classe 1 apresenta *precision* de 100%, mas *recall* de 66,67%, resultando em um F1 de 80%. Isso significa que, apesar de não haver falsos positivos (tudo o que foi predito como classe 1 de fato o é), alguns exemplos reais de classe 1 não foram detectados. Já a classe 2 fica em 66,67% tanto para *precision* quanto para *recall*, sugerindo um desempenho mais equilibrado, porém mais modesto, indicando que parte das instâncias dessa classe ainda é confundida. Por fim, a classe 3 exibe *recall* de 100% (todos os exemplos de classe 3 foram capturados), mas com 66,67% de *precision*, implicando em algumas classificações indevidas como classe 3. Essas variações reforçam a importância de, no aperfeiçoamento do modelo, observar cuidadosamente o trade-off entre *precision* e *recall* em cada classe, especialmente para garantir que nenhuma delas seja sub-representada ou confundida com frequência.

4. AVALIAÇÃO DO PREPPY

A motivação para aplicar o algoritmo PREPPY ao conjunto de dados *1985 Auto Imports* (SCHLIMMER, 1985) decorre das complexidades apresentadas por um *dataset* com atributos heterogêneos. Ao transformar critérios em atributos categóricos e definir preferências específicas que reflitam o risco veicular – com o objetivo de identificar, com precisão, veículos de menor risco e estabelecer relações entre categorias – evidencia-se a relevância de testar o algoritmo em cenários desafiadores. Importante ressaltar que a escolha deste conjunto específico, voltado para a avaliação do risco automotivo, serve apenas como exemplo ilustrativo. A estrutura geral do PREPPY foi concebida com uma capacidade generalista, permitindo sua aplicação em uma vasta gama de contextos, como avaliação de desempenho, classificação de investimentos e análise de riscos em setores diversos.

Ademais, ressalta-se que o PREPPY se destaca na área de engenharia de produção, demonstrando sua versatilidade ao ser aplicado na otimização de processos produtivos, no rigoroso controle de qualidade e na análise estratégica de riscos. Essa capacidade de adaptação evidencia seu potencial para apoiar a tomada de decisões em ambientes industriais desafiadores, ampliando significativamente o leque de aplicações do método. Cabe ainda destacar a contribuição da indução de regras de classificação na criação de um modelo representativo do processo decisório (SLOWINSKI, GRECO e MATARAZZO, 2015). Esse modelo capta padrões a partir de decisões previamente estabelecidas, permitindo a generalização das regras identificadas para um conjunto mais amplo de dados (GRECO et al., 2001). Assim, o algoritmo não apenas absorve as especificidades de um contexto particular, como também gera insights estratégicos valiosos para serem aplicados em outros cenários, reforçando sua eficácia e adaptabilidade em diversas áreas (BLASZCZYNSKI, SLOWINSKI, R e SZELAG, 2011; LEZANSKI e PILACINSKA, 2016).

Paralelamente, a comparação detalhada entre o PREPPY e o algoritmo DomLEM visa não somente avaliar a eficácia das regras induzidas, mas também analisar a robustez e a eficiência computacional dos métodos em cenários multicritério e multiclasse. Essa abordagem comparativa contribui para o aprimoramento dos métodos de classificação tanto em contextos reais como simulados, fortalecendo o entendimento sobre as vantagens e limitações inerentes a cada técnica.

4.1. Aplicação do PREPPY

Na aplicação do PREPPY ao conjunto *1985 Auto Imports* (SCHLIMMER, 1985), observam-se desafios significativos para a classificação do risco veicular. Com 205 instâncias e 26 atributos heterogêneos – distribuídos entre critérios e atributos categóricos –, o *dataset* requer um pré-processamento aprofundado, que inclui a transformação de variáveis e a definição precisa de preferências e relações entre categorias para representar adequadamente os fatores de risco. Esses desafios reforçam não só a adequação do algoritmo à complexidade dos dados, mas também a versatilidade do PREPPY, que está preparado para operar em diversas áreas onde a análise de dados heterogêneos e a indução de regras decisórias se fazem necessárias.

4.1.1. O Conjunto de Dados 1985 Auto Imports

O *1985 Auto Imports Database* (SCHLIMMER, 1985) consiste em 205 instâncias e 26 atributos, sendo 15 contínuos, 1 inteiro e 10 nominais. Tais atributos incluem características técnicas (por exemplo, Distância entre Eixos, Potência, Cilindrada) e categóricas (por exemplo, Fabricante (Credibilidade), Tipo de Combustível, Estilo de Carroceria) relacionadas ao risco veicular. Originalmente, o conjunto de dados apresenta valores ausentes em diversos atributos e exige um processo de limpeza para remoção ou imputação de dados incompletos.

Na prática, optou-se pela exclusão do atributo de Perdas Normalizadas pela alta quantidade de 41 valores em branco e alternativas com ao menos um valor ausente, resultando em 26 atributos e 193 casos após o pré-processamento. Também foi necessário transformar as variáveis numéricas de preço e altura em categorias não ordinais para simplificar a criação de regras, pois se constatou maior aderência às relações de desigualdade entre seus valores. Esse cenário impõe desafios devido à heterogeneidade entre atributos contínuos, categóricos e ordinais, demandando métodos capazes de lidar com diferentes naturezas de variáveis. Exemplo do conjunto disponibilizado no apêndice. A seguir, a Tabela 12 contém uma breve explicação de cada um dos atributos.

Tabela 12 – Descrição dos atributos

Atributo	Explicação
Número de Portas	Número de portas do veículo
Número de Cilindros	Quantidade de cilindros do motor

Distância entre Eixos	Distância entre os eixos dianteiro e traseiro, influenciando estabilidade e espaço interno.
Comprimento	Comprimento total do veículo.
Largura	Largura do veículo.
Peso sem Carga	Peso do veículo sem passageiros ou carga.
Cilindrada	Volume de deslocamento do motor.
Diâmetro Interno	Diâmetro interno de cada cilindro do motor.
Curso	Distância que o pistão percorre dentro do cilindro.
Taxa de Compressão	Relação entre o volume do cilindro antes e depois do curso de compressão, afetando desempenho e consumo.
Potência	Potência do motor.
RPM Máximo	Rotação máxima (em RPM) na qual o motor atinge sua potência máxima.
Consumo Urbano (MPG)	Consumo de combustível em milhas por galão na condução urbana.
Consumo Rodoviário (MPG)	Consumo de combustível em milhas por galão na condução em rodovias.
Fabricante (Credibilidade)	Fabricante ou marca do veículo onde é considerado sua credibilidade
Tipo de Combustível	Tipo de combustível.
Aspiração	Indica se o motor é aspirado ou turboalimentado.
Estilo de Carroceria	Formato da carroceria.
Tipo de Tração	Tipo de tração do veículo.
Localização do Motor	Localização do motor no veículo.
Tipo de Motor	Arquitetura interna do motor.
Sistema de Combustível	Sistema de alimentação de combustível.
Altura	Altura total do veículo.
Preço	Preço de mercado do veículo, em dólares (na época do conjunto de dados).

Fonte: O autor (2025)

Além dos procedimentos de pré-processamento e adaptação do algoritmo, foi necessário definir explicitamente as preferências e as relações entre categorias para orientar a indução das regras. Essas definições foram formuladas com o objetivo de priorizar a classe 6, que representa o menor risco, de forma que as configurações reflitam as preocupações inerentes a um carro de risco reduzido.

No que tange às preferências, os valores foram atribuídos considerando os aspectos que favorecem a segurança e a estabilidade do veículo. Assim, para os atributos numéricos e categóricos que influenciam a avaliação de risco, optou-se por classificá-los como

gain ou cost, de modo que a maximização dos ganhos (ou minimização dos custos) esteja alinhada com a obtenção de veículos de menor risco. Por exemplo, a atribuição de gain para atributos como Distância entre Eixos, Comprimento e Largura visa refletir a vantagem de dimensões que, de maneira holística, indicam maior robustez ou estabilidade, enquanto atributos que podem aumentar o risco foram configurados como cost. Especificamente, as preferências foram definidas na seguinte ordem: Número de Portas, Distância entre Eixos, Comprimento, Largura, Consumo Urbano (MPG) e Consumo Rodoviário (MPG) são configurados como gain; Número de Cilindros, Peso sem Carga, Cilindrada, Diâmetro Interno, Curso, Taxa de Compressão, Potência e RPM Máximo como cost; os demais atributos (Fabricante (Credibilidade), Tipo de Combustível, Aspiração, Estilo de Carroceria, Tipo de Tração, Localização do Motor, Tipo de Motor, Sistema de Combustível, Altura e Preço) são categóricos.

No que diz respeito às relações entre categorias, essas foram estabelecidas levando em conta a natureza ordinal dos atributos, quando esta estava claramente definida, e também com base em conhecimentos prévios acerca do setor automotivo. Foram definidas as relações a partir de uma escala tal como na Tabela 13. Cada escala é adequada ao número de categorias e das relações entre si em cada atributo categórico ordinal, consecutivamente gerando uma sequência de relações tal como na Tabela 14.

Tabela 13 – Definição da escala do atributo *Fabricante (Credibilidade)*

Categoria	Fabricante (Credibilidade)
Muito ruim	Alfa-Romeo, Porsche
Ruim	Audi, Mitsubishi, Volkswagen
Médio	Chevrolet, Dodge, Honda, Isuzu, Mercury, Nissan, Plymouth
Bom	Bmw, Honda, Jaguar, Mazda, Peugeot, Subaru
Muito bom	Mercedes-Benz, Toyota, Volvo

Fonte: O autor (2025)

Tabela 14 – Relações entre categorias

Atributo	Relações
Fabricante (Credibilidade)	Mercedes-Benz=Toyota=Volvo>Bmw=Honda=Jaguar=Mazda=Peugot=Subaru>Chevrolet=Dodge=Honda=Isuzu=Mercury=Nissan=Plymouth>Audi=Mitsubishi=Volkswagen>Alfa-Romeo=Porsche
Tipo de Combustível	Diesel>Gasolina
Aspiração	Std>Turbo

Estilo de Carroceria	<i>Wagon=Sedan>Hatchback>Hardtop=Convertible</i>
Tipo de Tração	Fwd>4Wd>Rwd
Localização do Motor	<i>Front>Rear</i>
Tipo de Motor	L>Ohc=Ohcf=Ohcv>Dohc=Dohcv> <i>Rotor</i>
Sistema de Combustível	1Bbl=2Bbl>4Bbl>Idi>Mfi=Mpfi>Spdi>Spfi

Fonte: O autor (2025)

Em que a relação de fabricantes reflete credibilidade, evidenciada pelo histórico de menor risco dos veículos da marca Volvo, por exemplo. É importante destacar que essas definições de preferências e relações entre categorias não são objetivas nem universais; elas foram estabelecidas a partir de uma avaliação holística da distribuição dos valores dos atributos ordinais nas classes, aliada a conhecimentos prévios sobre o comportamento de risco associado a cada valor.

Dessa forma, embora haja respaldo empírico e conceitual para tais escolhas, reconhece-se que elas podem ser revistas ou ajustadas à luz de novas evidências ou contextos de aplicação ou à subjetividade do tomador de decisão. Ressalto também que a divisão entre o conjunto de treinamento e *deployment* ocorreu de maneira aleatória. O PREPPY foi originalmente desenvolvido em um contexto de iniciação científica e, para esta aplicação, sofreu adaptações para computar o número extenso de atributos.

Em particular, configurou-se o algoritmo para induzir primeiro as regras dos atributos categóricos (ordinais ou não), devido à facilidade de processamento dessas variáveis por serem descritivas, e só então induzir as regras relativas aos critérios. Dessa forma, conseguiu-se lidar de maneira incremental com os dados heterogêneos. Em consequência, o método definiu regras para 9 dos 10 atributos categóricos disponíveis e para apenas 1 dos 14 critérios, o que evidenciou a alta demanda de recursos computacionais no tratamento de atributos contínuos de forma exaustiva.

4.1.2. Apresentação dos Resultados Experimentais

Inicialmente fora calculado o ganho de informação de cada um dos atributos para definir a priorização destes na permutação. Tomando como exemplo os que se destacam na Tabela 15.

Tabela 15 – Resultado dos cálculos de ganho de informação dos atributos

Nome	Fabricante (Credibilidade)	Distância entre Eixos	Peso sem Carga	Largura	Cilindrada	Comprimento
Score	1,1921	0,9673	0,7828	0,7782	0,7517	0,6825

Fonte: O autor (2025)

Ao final do processo de treinamento, sendo utilizado o conjunto de aproximação inferior, o PREPPY gerou um conjunto de regras de classificação de risco, agrupando as instâncias em seis classes (de 1 a 6, variando de maior para menor risco). As regras induzidas demonstram como o PREPPY consegue combinar atributos, como reparado na Tabela 16. Para cada classe, foram definidas combinações de atributos como Fabricante (Credibilidade), Tipo de Combustível, Aspiração, Estilo de Carroceria, Tipo de Tração, entre outros. Em especial, observou-se que:

- A classe 6, associada ao menor risco, apresentou regras bastante restritivas, envolvendo atributos como Fabricante (Credibilidade) (por exemplo, Volvo) e Sistema de Combustível (mpfi), com Altura categorizada como Alta;
- As classes intermediárias (2, 3, 4 e 5) incluíram atributos mais gerais, pois englobam maior diversidade de instâncias no conjunto de dados;
- Notou-se que a regra para a classe 2, embora tivesse *precision* elevada (0,89), alcançou *recall* de apenas 0,5 no treinamento, o que indica certa limitação na abrangência.

Tabela 16 – Regras induzidas

Atributo	Classe 6	Classe 5
<i>Distância entre Eixos</i>	≥ 10351	≥ 10121
<i>Fabricante (Credibilidade)</i>	<i>Volvo</i>	<i>Volvo or Toyota or Mercedes-Benz or Dodge or Mitsubishi</i>
<i>Tipo de Combustível</i>	<i>Gas</i>	<i>Gas or Diesel</i>
<i>Aspiração</i>	<i>Turbo</i>	<i>Std or Turbo</i>
<i>Estilo de Carroceria</i>	<i>Sedan</i>	<i>Sedan or Wagon or Hatchback</i>
<i>Tipo de Tração</i>	<i>Rwd</i>	<i>Rwd or Fwd</i>
<i>Localização do Motor</i>	<i>Front</i>	<i>Front</i>
<i>Tipo de Motor</i>	<i>Ohc</i>	<i>Ohc or Ohcv or Dohc</i>
<i>Sistema de Combustível</i>	<i>Mpfi</i>	<i>Mpfi or Idi or 2Bbl or Spdi</i>
<i>Altura</i>	<i>Alto</i>	<i>Médio or Alto or Baixo</i>

Atributo	Classe 4
Distância entre Eixos	≥ 9451
Fabricante (Credibilidade)	Toyota or Volvo or Peugeot or Subaru or Bmw or Mercedes-Benz or Mazda or Honda or Volkswagen or Dodge or Mitsubishi or Nissan or Jaguar
Tipo de Combustível	Gas or Diesel
Aspiração	Std or Turbo
Estilo de Carroceria	Sedan or Wagon or Hatchback
Tipo de Tração	Rwd or Fwd or 4Wd
Localização do Motor	Front
Tipo de Motor	Ohc or L or Ohcf or Ohcv or Dohc
Sistema de Combustível	Mpfi or 2Bbl or Idi or 1Bbl or Spdi
Altura	Médio or Alto or Baixo
Atributo	Classe 3
<i>Distância entre Eixos</i>	≥ 9131
<i>Fabricante (Credibilidade)</i>	<i>Toyota or Mazda or Nissan or Volvo or Peugeot or Subaru or Dodge or Bmw or Mercedes-Benz or Honda or Mitsubishi or Plymouth or Volkswagen or Audi or Jaguar or Alfa-Romeo</i>
<i>Tipo de Combustível</i>	<i>Gas or Diesel</i>
<i>Aspiração</i>	<i>Std or Turbo</i>
<i>Estilo de Carroceria</i>	<i>Sedan or Hatchback or Wagon</i>
<i>Tipo de Tração</i>	<i>Fwd or Rwd or 4Wd</i>
<i>Localização do Motor</i>	<i>Front</i>
<i>Tipo de Motor</i>	<i>Ohc or L or Ohcf or Ohcv or Dohc</i>
<i>Sistema de Combustível</i>	<i>Mpfi or 2Bbl or Idi or Spdi or 1Bbl</i>
<i>Altura</i>	<i>Médio or Baixo or Alto</i>
Atributo	Classe 2
<i>Distância entre Eixos</i>	≥ 8659
<i>Fabricante (Credibilidade)</i>	<i>Toyota or Mazda or Nissan or Subaru or Volvo or Peugeot or Bmw or Honda or Volkswagen or Dodge or Mitsubishi or Mercedes-Benz or Plymouth or Audi or Saab or Isuzu or Chevrolet or Jaguar or Alfa-Romeo</i>
<i>Tipo de Combustível</i>	<i>Gas or Diesel</i>

<i>Aspiração</i>	<i>Std or Turbo</i>
<i>Estilo de Carroceria</i>	<i>Sedan or Hatchback or Wagon or Convertible or Hardtop</i>
<i>Tipo de Tração</i>	<i>Fwd or Rwd or 4Wd</i>
<i>Localização do Motor</i>	<i>Front</i>
<i>Tipo de Motor</i>	<i>Ohc or L or Ohcf or Ohcv or Dohc</i>
<i>Sistema de Combustível</i>	<i>Mpfi or 2Bbl or Idi or 1Bbl or Spdi or Spfi</i>
<i>Altura</i>	<i>Médio or Baixo or Alto</i>

Fonte: O autor (2025)

As quais geraram a matriz de confusão do *predicted* com as classes originais e suas métricas, representada pela Tabelas 17, 18 e 19.

Tabela 17 – Matriz de confusão do previso com a classe original no treinamento

	Classe 6	Classe 5	Classe 4	Classe 3	Classe 2	Classe 1
Classe 6	1	0	0	0	0	0
Classe 5	0	13	1	0	0	0
Classe 4	0	1	30	0	0	0
Classe 3	0	0	4	20	0	0
Classe 2	0	0	5	3	8	0
Classe 1	0	1	3	1	1	5

Fonte: O autor (2025)

Tabela 18 – Métricas gerais do treinamento

<i>Metric</i>	<i>Value</i>
<i>Accuracy</i>	79,38%
<i>Precision (Macro)</i>	88,11%
<i>Precision (Micro)</i>	79,38%
<i>Precision (Weighted)</i>	82,46%
<i>Recall (Macro)</i>	78,07%
<i>Recall (Micro)</i>	79,38%
<i>Recall (Weighted)</i>	79,38%
<i>F1-Score (Macro)</i>	80,09%
<i>F1-Score (Micro)</i>	79,38%
<i>F1-Score (Weighted)</i>	78,15%
<i>Matthews Correlation Coefficient (MCC)</i>	73,91%
<i>Cohen's Kappa</i>	72,64%

Fonte: O autor (2025)

Tabela 19 – Métricas por classe do treinamento

Métrica	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Classe 1	100,00%	45,45%	62,50%
Classe 2	88,89%	50,00%	64,00%
Classe 3	83,33%	83,33%	83,33%

Classe 4	69,77%	96,77%	81,08%
Classe 5	86,67%	92,86%	89,66%
Classe 6	100,00%	45,45%	62,50%

Fonte: O autor (2025)

A análise da matriz de confusão e das métricas de desempenho revela que o modelo apresenta um desempenho global robusto, com um *accuracy* próximo a 79% e um *Matthews Correlation Coefficient* em torno de 74%, evidenciando boa correlação entre as predições e as classes originais. O sistema se destaca na identificação dos extremos: veículos de menor risco (classe 6) foram classificados com perfeição, enquanto, apesar de a classe 1 ter alcançado 100% de *precision*, seu *recall* significativamente inferior indica que muitos casos dessa classe não foram identificados. Por outro lado, as classes intermediárias exibem um equilíbrio moderado entre *precision* e *recall*, com índices geralmente acima de 80%, mas ainda sofrem de certa ambiguidade na diferenciação, o que se reflete na ocorrência de erros na atribuição entre as classes 2, 3 e 4. Em síntese, o modelo demonstra excelência na classificação de veículos de risco extremo, enquanto a distinção fina entre níveis intermediários aponta para oportunidades de aperfeiçoamento.

Para avaliar a generalização do PREPPY, as regras induzidas foram aplicadas em um cenário de *deployment*, onde os dados são distintos ou representam novas observações. Obteve-se os resultados e suas métricas, demonstrados pelas Tabelas 21 e 22.

Tabela 21 – Matriz de confusão do previso com a classe original no *deployment*

	Classe 6	Classe 5	Classe 4	Classe 3	Classe 2	Classe 1
Classe 6	2	0	0	0	0	0
Classe 5	0	7	0	1	0	0
Classe 4	0	0	28	0	3	1
Classe 3	0	0	6	16	2	3
Classe 2	0	0	5	6	4	0
Classe 1	0	1	0	1	4	6

Fonte: O autor (2025)

Tabela 22 – Métricas gerais do *deployment*

Metric	Value
<i>Accuracy</i>	65,63%
<i>Precision (Macro)</i>	69,46%
<i>Precision (Micro)</i>	65,63%
<i>Precision (Weighted)</i>	64,36%
<i>Recall (Macro)</i>	68,49%
<i>Recall (Micro)</i>	65,63%
<i>Recall (Weighted)</i>	65,63%
<i>F1-Score (Macro)</i>	68,71%

<i>F1-Score (Micro)</i>	65,63%
<i>F1-Score (Weighted)</i>	64,60%
<i>Matthews Correlation Coefficient (MCC)</i>	54,60%
<i>Cohen's Kappa</i>	54,33%

Fonte: O autor (2025)

abela 23 – Métricas por classe do *deployment*

Métrica	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Classe 1	60,00%	50,00%	54,55%
Classe 2	30,77%	26,67%	28,57%
Classe 3	66,67%	59,26%	62,75%
Classe 4	71,79%	87,50%	78,87%
Classe 5	87,50%	87,50%	87,50%
Classe 6	100,00%	100,00%	100,00%

Fonte: O autor (2025)

Observa-se um decréscimo em todos os indicadores, com *accuracy* passando de 0,79 para 0,65. Ainda assim, a classe 6 manteve *precision*, *recall* e *f1-score* em 1, mostrando que o método continua eficaz para identificar instâncias de maior risco. Por outro lado, a classe 2 sofreu queda expressiva em *precision* (0,3077) e *recall* (0,2667), evidenciando maior dificuldade na generalização para esse grupo específico. A matriz de confusão no *deployment* reforça a concentração de acertos em classes de maior risco e certa confusão entre classes próximas (por exemplo, 2, 3 e 4).

4.1.3. Análise e Discussão dos Resultados

A análise dos resultados experimentais obtidos com o PREPPY evidencia, de maneira geral, uma performance consistente na classificação de veículos quanto ao risco, especialmente na identificação em melhores classes do espectro de risco. A metodologia adotada – que combina regras derivadas de critérios e atributos categóricos – demonstrou eficiência ao lidar com a heterogeneidade dos dados presentes no conjunto 1985 *Auto Imports* (SCHLIMMER, 1985), mesmo diante de desafios inerentes à transformação de variáveis contínuas em categorias não ordinais, à exclusão de atributos com elevada incidência de valores ausentes e problemas de capacidade computacional.

Os resultados têm implicações práticas importantes para o setor automotivo. A alta *precision* e *recall* na classe 6 demonstram a eficácia do método em identificar veículos com menor risco, aspecto valioso para seguros e inspeções veiculares. Em contrapartida, o desempenho insatisfatório na classe 2, evidenciado pela baixa abrangência e *precision*, alerta para a necessidade de cautela quando se busca discriminação mais refinada entre níveis de risco.

Assim, embora o PREPPY se destaque na classificação dos casos extremos, aprimoramentos são necessários para melhorar sua generalização em cenários com maior ambiguidade.

A queda expressiva nas métricas da classe 2 no *deployment* pode ser atribuída à heterogeneidade dos dados, ao pré-processamento que incluiu a exclusão do atributo *Perdas Normalizadas* e de casos com dados ausentes – o que, embora tenha tornado o conjunto mais homogêneo, reduziu a variabilidade necessária para capturar nuances entre classes – e à definição de preferências, que favoreceu a identificação dos extremos, mas pode não refletir adequadamente as interações complexas presentes nas classes de risco intermediário.

Ao comparar os resultados com estudos da literatura, observa-se que métodos de classificação baseados em regras tendem a identificar os extremos – veículos com riscos muito altos ou muito baixos – de forma mais assertiva, enquanto a delimitação das classes intermediárias permanece desafiadora. Essa tendência reflete o objetivo comum dos sistemas de aprendizado, que é encontrar uma descrição discriminante do conceito (GRZYMALA-BUSSE, 1992), e a própria natureza da teoria dos *rough sets*, que calcula as aproximações inferior e superior dos conceitos, evidenciando as dificuldades inerentes à definição precisa dos intervalos intermediários (PAWLAK, 1991; ZIARKO, 1993).

A integração de conhecimento prévio, como as relações entre marcas e atributos técnicos, pode aprimorar a interpretabilidade dos modelos, embora frequentemente em detrimento da generalização em cenários com alta variabilidade. Abordagens que buscam um equilíbrio entre *precision* e generalidade, como a indução de regras de consistência variável (BLASZCZYNSKI, SLOWINSKI, R e SZELAG, 2011), são valorizadas por sua capacidade de oferecer explicações claras para as classificações (GRECO et al., 2001). Nesse contexto, a estratégia baseada em dominância (DRSA) adotada pelo PREPPY apresenta desempenho diferenciado nas classes superiores, mas pode ser afetada por desafios como a especificidade do conhecimento incorporado e o risco de *overfitting*, o que ressalta a necessidade de validação contínua e adaptação aos novos dados (GRZYMALA-BUSSE, 1992; KATO, ITSUNO e SAEKI, 2017). A utilidade dessas abordagens para a identificação eficaz de casos de alto ou mais baixo risco em ambientes práticos é corroborada pela literatura (DO COUTO e RANGEL, 2024).

Os resultados do PREPPY evidenciam um método promissor para a classificação de risco, destacando sua eficácia na identificação de veículos de menor risco (especialmente na classe 6) e sua capacidade de lidar com a heterogeneidade dos dados, embora a generalização

para classes inferiores ainda presente desafios. Isso ressalta a necessidade de ajustes metodológicos, seja na definição de preferências ou no tratamento de variáveis contínuas, e reforça a importância de abordagens híbridas que combinem conhecimento prévio com técnicas adaptativas, alinhando o algoritmo às melhores práticas da literatura e às exigências de cenários operacionais reais, onde a segurança e a *precision* na avaliação de risco são essenciais.

4.2. Avaliação Pela Comparação Entre PREPPY e DomLEM

Na seção de Comparação, foi realizada uma comparação rigorosa entre a nova abordagem PREPPY e o legado DomLEM, com grande preocupação em validar se as inovações do novo algoritmo superam o método tradicional nas instâncias consideradas. Utilizando *datasets* sintéticos que simulam cenários reais, foram avaliadas objetivamente métricas de desempenho, robustez e eficiência computacional, garantindo uma análise crítica e fundamentada dos resultados.

4.2.1. Os Conjuntos de Dados Sintéticos

Os *datasets* sintéticos utilizados neste estudo foram fundamentados a partir de fontes reais, com o objetivo de simular cenários multicritério e multiclasse de forma realista. Esses conjuntos apresentam variáveis numéricas contínuas e discretas, refletindo contextos práticos como eficiência energética, desempenho de veículos elétricos, análise de aplicativos móveis e avaliação de docentes. Como por exemplo na Tabela 25.

Tabela 25 – Exemplo do conjunto de três atributos e cem alternativas

Alternativas	Consumo Energia (kWh/m ² /ano)	Conforto Térmico (ICT)	Custo Implementação (R\$/m ²)	Classe
1	147,5357	0,5491	420,13	2
2	157,7997	0,0232	673,24	1
3	84,6956	0,6139	294,01	3
4	18,2812	0,8445	26,51	4
5	126,2378	0,4921	358,95	2

Fonte: O autor (2025)

A decisão de simular os dados não só possibilita o controle das distribuições e a simulação de desafios típicos (como outliers e classes adjacentes), mas também permite reproduzir pequenas inconsistências encontradas em ambientes reais. Essa estratégia assegura que os atributos dos conjuntos sejam cuidadosamente adequados, o que responde à necessidade de simular desafios práticos e reforça a robustez dos algoritmos, para ambos algoritmos PREPPY e DomLEM. Além disso, o balanceamento dos dados evita a dominância de uma única

classe, garantindo que a divisão entre treinamento e *deployment* seja representativa e contribuindo para a validação da capacidade de generalização dos modelos.

Dessa forma, os conjuntos de dados são compostos por variáveis essenciais para a análise (como ganhos ou custos) e organizados de maneira que tanto o treinamento quanto o *deployment* reflitam fielmente os cenários reais, respondendo à questão de como a estrutura dos *datasets* impacta a robustez e a eficácia dos algoritmos. Ressalto também que os conjuntos foram separados em *training* e *deployment* de maneira aleatória, com ajustes precisos para que haja uma boa distribuição das classes entre os dois conjuntos

4.2.2. Procedimentos de Comparação

A metodologia de comparação adotada visa comparar, de forma objetiva, os algoritmos PREPPY (implementado em Python) e DomLEM (disponível no JMAF, baseado em Java), ambos fundamentados na *Rough Set Theory* (RST) e na variante baseada em dominância (DRSA). Para assegurar uma avaliação consistente, os dados foram divididos em dois subconjuntos: o conjunto de treinamento, onde são induzidas as regras de decisão, e o conjunto de *deployment*, utilizado para testar a generalização dos modelos em dados inéditos.

Os experimentos foram realizados variando o número de critérios (3, 5, 7 e 9) e alternativas (100, 200, 300 e 400), permitindo observar como suas variações influenciam tanto a qualidade das regras induzidas quanto o desempenho geral dos algoritmos. Essas variações demonstram que, embora a inclusão de mais critérios possa enriquecer o modelo com informações adicionais, ela também pode introduzir ruído e aumentar a complexidade das decisões, o que impacta diretamente as métricas de avaliação.

Para medir o desempenho, foram adotadas métricas essenciais como acurácia, *precision*, *recall*, *F1-Score* (*macro* e *micro*), Coeficiente de Correlação de Matthews (MCC), Kappa de Cohen e, quando pertinente, o tempo de execução. A divisão dos dados entre treinamento e *deployment* não apenas assegura a validade dos testes – permitindo comparar o desempenho em ambientes conhecidos e inéditos – mas também facilita a identificação de eventuais discrepâncias e a avaliação da robustez dos algoritmos. Por exemplo, enquanto o PREPPY tendeu a induzir um conjunto reduzido e interpretável de regras, o DomLEM gerou um maior número de regras, muitas vezes com conjunções de condições que levaram a um risco mais elevado de *overfitting* quando aplicado aos dados do *deployment*.

Ademais, o tempo de execução de cada método foi registrado desde o início da indução das regras até a aplicação no conjunto de *deployment*, o que possibilita uma comparação direta da eficiência computacional entre as implementações em Python e Java. Essa abordagem detalhada assegura que todas as variáveis – desde a complexidade dos dados até os aspectos computacionais – sejam levadas em conta na análise comparativa dos métodos, integrando as respostas às questões sobre métricas, divisão dos dados, variação dos atributos e eficiência.

Além dos procedimentos já descritos, foram desenvolvidos scripts em Python para converter o arquivo Excel padronizado (semelhante ao do PREPPY) em formato TXT compatível com o JMAF, bem como para ler o output do DomLEM e reorganizá-lo em um arquivo Excel unificado. Essa automação inclui a geração de matrizes de confusão, o cálculo das mesmas métricas de avaliação utilizadas pelo PREPPY (acurácia, *precision*, *recall*, *F1-Score macro e micro*, MCC e Kappa de Cohen) e a criação de abas específicas para regras induzidas, exemplos utilizados e resultados de classificação. Assim, garante-se uma comparação uniforme entre os dois algoritmos, integrando as etapas de preparação dos dados, indução de regras e análise de desempenho em um único fluxo de trabalho.

4.2.3. Apresentação e Discussão das Métricas

A habilidade de generalização das regras ocupa um lugar central na análise de comparação. Avaliar as métricas na fase de *deployment* permite medir como as regras induzidas pelos algoritmos mantêm seu desempenho em dados inéditos.

No cenário com três critérios, o PREPPY apresentou resultados consistentes e superiores ao DomLEM. Sua acurácia geral foi de 84,00%, com métricas macro igualmente robustas (*precision*: 83,79%; *recall*: 86,69%; F1: 83,76%; MCC: 79,06%; Kappa: 78,28%). Destaca-se a Classe 4 com elevado *recall*, sugerindo regras mais inclusivas ou eficazes em novos dados. Já a Classe 1 exibiu uma *precision* relativamente menor, mas sem comprometer a acurácia geral. O DomLEM alcançou uma acurácia mais modesta (74,00%) e apresentou flutuações maiores entre classes. Por exemplo, o alto *recall* na Classe 1 contrasta com a baixa *precision* na Classe 4, indicando potenciais problemas de generalização ou confusão entre categorias. Assim, o PREPPY demonstrou melhor estabilidade e equilíbrio em relação ao DomLEM.

No cenário com cinco critérios, o desempenho do PREPPY permaneceu satisfatório, embora com uma leve queda (acurácia de 68,00%; *precision* macro: 76,11%; *recall* macro: 68,74%; F1 macro: 68,87%; MCC: 57,95%; Kappa: 56,24%). A Classe 4 teve novamente destaque, com aumentos notáveis no *recall* e no F1, indicando maior eficácia das regras geradas. As demais classes mantiveram estabilidade nos resultados. Em contraste, o DomLEM registrou desempenho inferior (acurácia: 49,00%; *precision* macro: 66,12%; *recall* macro: 46,52%; F1 macro: 43,66%; MCC: 31,67%; Kappa: 28,83%). Apesar da boa *precision* da Classe 4, seu baixo *recall* revelou regras restritivas demais, prejudicando significativamente o resultado geral.

Com sete critérios, o PREPPY recuperou parcialmente suas métricas anteriores, atingindo acurácia de 74,67%, *precision* macro de 76,68%, *recall* macro de 75,30% e F1 macro de 75,78%, com MCC (66,06%) e Kappa (65,94%) satisfatórios. A Classe 4 se destacou novamente com resultados excepcionais, superiores a 100% em *precision* e *recall*, enquanto as demais classes tiveram desempenho sólido. Já o DomLEM apresentou resultados inferiores, especialmente na Classe 3, cuja queda acentuada indicou dificuldades claras na generalização das regras (acurácia: 59,33%; *precision* macro: 66,23%; *recall* macro: 60,57%; F1 macro: 61,18%; MCC: 46,79%; Kappa: 45,24%).

No cenário de nove critérios, o PREPPY alcançou métricas ligeiramente superiores ao conjunto anterior (acurácia: 75,00%; *precision* macro: 77,23%; *recall* macro: 75,40%; F1 macro: 75,89%; MCC: 66,85%; Kappa: 66,52%). Destaca-se novamente o *recall* elevado da Classe 1, reforçando a boa adaptação do método ao conjunto de testes. O DomLEM, por outro lado, teve desempenho mais moderado, com acurácia de 68,00%, *precision* macro de 73,07%, *recall* macro de 67,92% e F1 macro de 69,25%, além de MCC (57,58%) e Kappa (56,91%) inferiores ao PREPPY. Destacou-se a Classe 1, que teve um *recall* elevado, sugerindo maior abrangência das regras, enquanto as demais classes apresentaram resultados apenas medianos.

Portanto, o PREPPY demonstrou consistentemente maior estabilidade, menor oscilação entre classes e melhor desempenho geral na generalização para dados inéditos em comparação ao DomLEM. Isso se deve à estrutura flexível das suas regras, com condições numéricas adaptáveis, permitindo uma cobertura efetiva de várias classes mesmo quando certos critérios não são determinantes. Por outro lado, o DomLEM, ao criar regras mais específicas com conjunções múltiplas, mostrou-se mais propenso ao *overfitting*, resultando em queda no

desempenho fora do conjunto de treino, especialmente em cenários com maior número de critérios. Entretanto, é importante salientar que o PREPPY pode ser mais lento devido à complexidade adicional das permutações necessárias no processamento das regras. Ademais, não se pode afirmar que o PREPPY seja sempre superior; em contextos específicos, como problemas com muitos atributos, o DomLEM pode apresentar vantagens por sua abordagem mais direta. Essas considerações ressaltam as vantagens contextuais do PREPPY, com a ressalva da possibilidade de cenários particulares em que o DomLEM seja mais eficiente.

5. CONCLUSÃO

5.1. Síntese dos Principais Resultados e Contribuições

O presente trabalho teve como objetivo principal desenvolver, introduzir e avaliar o algoritmo PREPPY para a indução de regras de classificação. Para sua avaliação, decorreu por princípio a classificação de risco em automóveis, seguido da realização de comparações com o DomLEM. Durante o estudo, o PREPPY foi desenvolvido para ser compatível com atributos diversos e explorar combinações destes para indução incremental de regras equilibradas entre *recall* e *precision*. Tendo também como preocupação a geração regras concisas, interpretáveis e generalizáveis.

Além do aspecto técnico, o desenvolvimento do PREPPY incorporou conceitos fundamentais da engenharia de produção, enfatizando a importância da tomada de decisão baseada em dados e a otimização dos processos decisórios em organizações. Essa abordagem integra práticas como a automatização e a integração com sistemas de informação, fundamentais para o apoio na gestão e na resolução de problemas de multicritério. A aplicação desses conceitos permite não apenas a melhoria da eficiência dos processos decisórios, mas também a criação de modelos que se adaptam a diferentes contextos.

Para avaliação da sua eficácia, fora primeiramente utilizado o conjunto de dados 1985 *Auto Imports* (SCHLIMMER, 1985), onde evidenciou-se sua capacidade de identificar corretamente as classes superiores – uma característica que se mostra fundamental não somente para o setor de seguros e inspeções veiculares, mas também para práticas de monitoramento e controle de qualidade. A avaliação seguiu com a realização de comparações entre o PREPPY e DomLEM utilizando uma gama de conjuntos de dados sintéticos, as quais o PREPPY se mostrou superior em todas as instâncias. Entre os resultados mais relevantes, destacam-se:

- Desempenho superior ao DomLEM: Na comparação realizada com conjuntos de dados sintéticos variando em quantidade de atributos e número de alternativas, o PREPPY se mostrou superior não só em todas as instâncias, mas em quase todas as métricas calculadas. Representando um avanço significativo, mas que requer maiores testes;
- Alto desempenho em classes superiores: O algoritmo demonstrou alta *precision* e *recall* na identificação de veículos com menor risco, evidenciando sua eficácia na delimitação de classes críticas.

- **Contribuições Teóricas e Práticas:** Em nível teórico, o trabalho promove os fundamentos do DRSA ao oferecer uma nova alternativa que se distingue das demais na indução de regras. Enquanto, na prática, também possibilita a extração de regras interpretáveis e robustas que podem ser utilizadas em cenários decisórios complexos.

Assim, o PREPPY contribui de forma significativa para a área de decisão multicritério – e, em particular, para a aplicação dos conceitos de RST – ambas especialmente pertinentes na engenharia de produção. Ele também oferece uma abordagem inovadora que alia rigor matemático à adaptabilidade para dados diversos, integrando técnicas de ganho de informação e permutação de critérios para promover uma estrutura decisória fundamentada em dados, mais ágil e assertiva, capaz de enfrentar os desafios de ambientes complexos e de múltiplos critérios.

5.2. Limitações do Estudo

Embora os resultados obtenham avanços expressivos, o estudo apresentou limitações que devem ser consideradas para futuras pesquisas, inclusive sob a ótica dos desafios enfrentados em ambientes de produção:

- **Complexidade Computacional:** O processo exaustivo de exploração das permutações dos critérios para o ajuste dos parâmetros demandou um elevado tempo de processamento. Essa limitação pode comprometer a escalabilidade do algoritmo quando aplicado a bases de dados maiores ou com um número ampliado de atributos, cenário comum em indústrias com grandes volumes de informação.
- **Sensibilidade a Dados e Pré-processamento:** A necessidade de transformar critérios em categorias não ordinais e a exclusão de um atributo com valores ausentes reduziram a variabilidade dos dados, o que pode limitar a generalização dos resultados. Além disso, a definição das preferências e das relações entre categorias, embora fundamentadas em conhecimento prévio, pode não captar todas as nuances presentes em contextos de alta heterogeneidade.

Essas limitações indicam que, apesar dos resultados promissores, os desafios na definição de regras para classes inferiores e a necessidade de maior eficiência computacional podem afetar a generalização dos resultados quando aplicados a outros contextos ou bases de dados mais complexas.

5.3. Sugestões para Trabalhos Futuros e Considerações Finais

Para ampliar a aplicabilidade e a robustez do PREPPY, bem como integrar ainda mais as práticas da Engenharia de Produção, sugerem-se as seguintes estratégias:

- Otimização computacional: Desenvolver cálculos vetorizados mais eficientes, implementar computação paralela e promover um maior controle de memória. Tais iniciativas reduzirão o tempo de processamento sem comprometer a *precision* do modelo, refletindo a busca contínua por eficiência, característica central na Engenharia de Produção.
- Testes com novos conjuntos de dados: Avaliar o algoritmo em outros conjuntos de dados – sejam eles reais ou sintéticos – para validar sua robustez e adaptabilidade. Essa abordagem permitirá ajustar as definições de preferências e relações entre categorias de acordo com os distintos domínios de aplicação, semelhante à adaptação de processos produtivos a diferentes cenários industriais.
- Refatoração do código segundo princípios modernos de arquitetura de software: Adotar filosofias como a separação de responsabilidades e padrões de projeto de programação, assegurando a modularidade, escalabilidade e a manutenção facilitada, por meio de pipelines de integração/entrega contínua, em linha com as práticas modernas de gestão de processos na Engenharia de Produção.
- Escalabilidade e acesso em nuvem: Hospedar a aplicação em plataformas de nuvem, garantindo escalabilidade, disponibilidade contínua e controle total de versionamento, além de recursos de processamento compatíveis com bases de dados ampliadas e um maior número de usuários simultâneos – uma estratégia bastante utilizada em ambientes industriais para assegurar a continuidade operacional.
- Tratamento de inconsistências: Introduzir mecanismos de tratamento de inconsistências por meio de conceitos de consistência variável, flexibilizando os critérios rígidos dos modelos *Rough Sets*. Essa abordagem permitirá considerar objetos com alta, mas não perfeita, evidência de pertencimento, ampliando a aplicabilidade do modelo em contextos com dados imperfeitos, frequentemente observados em operações industriais.

Portanto, o trabalho evidencia a relevância do PREPPY como uma ferramenta inovadora de apoio à decisão multicritério, especialmente em contextos que exigem a integração de dados heterogêneos e a otimização dos processos decisórios. Ao incorporar fundamentos da engenharia de produção – como a tomada de decisão orientada por dados, a

automação dos fluxos de trabalho e a integração com sistemas de informação –, o modelo demonstra aplicabilidade em ambientes que demandam eficiência e uma gestão assertiva. Os aprendizados obtidos durante o estudo – que englobam desde os fundamentos de RST e a arquitetura de software até a gestão de projetos de computação, a pesquisa do estado da arte e a exploração criativa para inovações – apontam para um caminho promissor de aprimoramento contínuo, com potencial para influenciar futuras pesquisas e aplicações práticas tanto na indústria quanto na academia. Dessa forma, com os ajustes sugeridos, o PREPPY tem o potencial de se consolidar como uma alternativa robusta em sistemas de apoio à decisão, contribuindo decisivamente para a evolução dos processos decisórios e para a melhoria global da gestão organizacional.

REFERÊNCIAS

BLASZCZYNSKI, J.; GRECO, S.; MATARAZZO, B.; SZELAĞ, M. Dominance-based rough set approach, basic ideas and main trends. arXiv, 2022.

BŁASZCZYŃSKI, J.; GRECO, S.; SŁOWIŃSKI, R.; SZELAĞ, M. On Variable Consistency Dominance-Based Rough Set Approaches. In: INTERNATIONAL CONFERENCE ON ROUGH SETS AND CURRENT TRENDS IN COMPUTING (RSCTC), 2006, Kobe. Anais. Berlin: Springer-Verlag, 2006. p. 191–202. (Lecture Notes in Artificial Intelligence, v. 4259).

BLASZCZYNSKI, J.; SLOWINSKI, R.; SZELAG, M. Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences*, v. 181, p. 987–1002, 2011.

BLASZCZYNSKI, J.; SLOWINSKI, R.; SZELAG, M. Monotonic variable consistency rough set approaches. *International Journal of Approximate Reasoning*, v. 50, n. 7, p. 979–999, 2009.

CHAKHAR, S.; SAAD, I. Dominance-based rough set approach for groups in multicriteria classification problems. *Decision Support Systems*, v. 54, p. 372–380, 2012.

CHAN, Y.; GRZYMALA-BUSSE, J. A rough set approach to rule extraction. *International Journal of Intelligent Systems*, v. 6, n. 5, p. 429–441, 1991.

DO COUTO, A.; RANGEL, L. A new application with conditional risk matrix and decision rules: an assessment of the sustainability and performance electric power company. *Foundations of Computing and Decision Sciences*, Warsaw, v. 49, n. 3, p. 233-260, 2024.

FAN, T.-F.; LIAU, C.-J.; LIU, D.-R. Variable consistency and variable precision models for dominance-based fuzzy rough set analysis of possibilistic information systems. *International Journal of General Systems*, v. 42, n. 6, p. 659–686, 2013.

GAETA, A.; LOIA, V.; LOMASTO, L.; ORCIUOLI, F. A novel approach based on rough set theory for analyzing information disorder. *Appl. Intell.*, v. 53, p. 15993–16014, 2023.

GIL-HERRERA, E.; ADEN-BUIE, G.; YALCIN, A.; TSALATSANIS, A.; BARNES, L. E.; DJULBEGOVIC, B. Rough set theory based prognostic classification models for hospice referral. *BMC Medical Informatics and Decision Making*, v. 15, p. 98, 2015.

GONG, Z.; LI, W. Multi-scale variable precision covering rough sets and its applications. *Applied Intelligence*, 2023. Aceito em 15 nov. 2023; publicado em 1 dez. 2023.

GOWIN, E.; JANUSZKIEWICZ-LEWANDOWSKA, D.; SLOWINSKI, R.; BLASZCZYNSKI, J.; MICHALAK, M.; WYSOCKI, J. With a little help from a computer: discriminating between bacterial and viral meningitis based on dominance-based rough set approach analysis. *Medicine*, v. 96, n. 32, p. e7635, 2017.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; STEFANOWSKI, J. An algorithm for induction of decision rules consistent with the dominance principle. In: *Rough Sets and Current Trends in Computing*, 2000. p. 304–313, 2000.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; STEFANOWSKI, J. Variable consistency model of dominance-based rough sets approach. In: *Rough Sets and Current Trends in Computing*, 2001. p. 170–181, 2001.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*, v. 129, n. 1, p. 1–47, 2001.

GRZYMALA-BUSSE, J. W. LERS-A system for learning from examples based on rough sets. In: *Intelligent Decision Support*, 1992. Dordrecht: Kluwer Academic Publishers, 1992. p. 3–18.

GRZYMALA-BUSSE, J. W. A new version of the rule induction system LERS. *Fundamenta Informaticae*, v. 31, n. 1, p. 27–39, 1997.

KARAMI, J.; GORSEVSKI, P. V.; JANKOWSKI, P. Water quality analysis using a variable consistency dominance-based rough set approach. *Computers, Environment and Urban Systems*, v. 43, p. 25–33, 2014.

KATO, Y.; ITSUNO, T.; SAEKI, T. Proposal of dominance-based rough set approach by STRIM and its applied example. In: *Rough Sets. International Joint Conference*. Springer. p. 418–431, 2017.

KOMOROWSKI, J.; POLKOWSKI, L.; SKOWRON, A. *Rough sets: a tutorial*. Dordrecht: Kluwer Academic Publishers, 1999.

KUSUNOKI, Yoshifumi; BŁASZCZYŃSKI, Jerzy; INUIGUCHI, Masahiro; SŁOWIŃSKI, Roman. Empirical risk minimization for dominance-based rough set approaches. *Information Sciences*, v. 567, p. 395–417, 2021.

LEZANSKI, P.; PILACINSKA, M. The dominance-based rough set approach to cylindrical plunge grinding process diagnosis. *Journal of Intelligent Manufacturing*, 2016.

LI, W.; XUE, X.; XU, W.; ZHAN, T.; FAN, B. Double-quantitative variable consistency dominance-based rough set approach. *International Journal of Approximate Reasoning*, v. 124, p. 1–26, 2020.

MUNKRES, JAMES. *Topology*. 2. ed. Upper Saddle River, NJ: Prentice Hill, 2000.

OGANGE, L. Hybridizing grid partitioning, rough set theory, and feature selection for fuzzy rule generation in dataset classification. *International Journal of Enterprise Modelling*, v. 17, n. 2, p. 79-87, 2023.

OPITZ, Juri. *A closer look at classification evaluation metrics and a critical reflection of common evaluation practice*. *Transactions of the Association for Computational Linguistics*, v. 12, p. 820–836, 2024.

PAWLAK, Z. Rough sets. *International Journal of Computer & Information Sciences*, v. 11, n. 5, p. 341–356, 1982.

PAWLAK, Z. *Rough sets: theoretical aspects of reasoning about data*. Dordrecht: Kluwer Academic Publishers, 1991.

PAWLAK, Z.; SLOWINSKI, R. Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research*, v. 72, n. 3, p. 443–459, 1994.

PAWLAK, Z.; SKOWRON, A. Rough membership functions. In: *Advances in the Dempster-Shafer Theory of Evidence*, 1994. New York: John Wiley & Sons, Inc. p. 251–271, 1994.

PAWLAK, Z.; GRZYMALA-BUSSE, J.; SLOWINSKI, R.; ZIARKO, W. Rough sets. *Communications of the ACM*, v. 38, n. 11, p. 88–95, 1995.

PAWLAK, Z.; SKOWRON, A. Rudiments of rough sets. *Information Sciences*, v. 177, n. 1, p. 3–27, 2007.

PAWLAK, Z. Learning from examples: the application of rough sets to the analysis of data. [S.l.]: Physica-Verlag, 1995.

SCHLIMMER, Jeffrey. *Automobile*. 1985. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5B01C>. Acesso em: 9 abr. 2025.

SHI, Z.; LI, L.; XIE, S.; XIE, J. The variable precision fuzzy rough set based on overlap and grouping functions with double weight method to MADM. *Applied Intelligence*, v. 54, p. 7696–7715, 2024.

SKOWRON, A. Abordagens baseadas em rough sets para mineração de dados. In: *Rough Set Approaches to Data Mining*, 2007.

SKOWRON, A.; DUTTA, S. Rough sets: past, present, and future. *Natural Computing*, v. 17, p. 855–876, 2018.

SLOWINSKI, R.; GRECO, S.; MATARAZZO, B. Rough set methodology for decision aiding. In: *Springer Handbook of Computational Intelligence*, 2015. Springer, 2015. p. 349–370.

SZELAG, M.; GRECO, S.; SLOWINSKI, R. Variable consistency dominance-based rough set approach to preference learning in multicriteria ranking. *Information Sciences*, v. 277, p. 525–552, 2014.

TOPAL, A.; GULER BAYAZIT, N.; UCAN, Y. A method to handle the missing values in multi-criteria sorting problems based on dominance rough sets. *Mathematics*, v. 12, n. 18, p. 2944, 2024.

WHITESITT, J. ELDON. *Boolean algebra and its applications*. New York: Dover Publications, 1995.

ZIARKO, W. Variable precision rough set model. *Journal of Computer and System Sciences*, v. 46, n. 1, p. 39–59, 1993.

APÊNDICE

Tabela 26 – Exemplos do conjunto de 1985 *Auto Imports*

<i>Alternativas</i>	<i>Fabricante (Credibilidade)</i>	<i>Tipo de Combustível</i>	<i>Aspiração</i>	<i>Estilo de Carroceria</i>	<i>Tipo de Tração</i>	<i>Localização do Motor</i>	<i>Tipo de Motor</i>	<i>Sistema de Combustível</i>	<i>Altura</i>
1	Bmw	Gas	Std	Sedan	Rwd	Front	Ohc	Mpfi	Médio
2	Audi	Gas	Std	Wagon	Fwd	Front	Ohc	Mpfi	Médio
3	Subaru	Gas	Std	Sedan	Fwd	Front	Ohcf	Mpfi	Baixo
4	Nissan	Gas	Std	Sedan	Fwd	Front	Ohc	2Bbl	Médio
5	Honda	Gas	Std	Hatchback	Fwd	Front	Ohc	1Bbl	Baixo
<i>Preço</i>	<i>Número de Portas</i>	<i>Número de Cilindros</i>	<i>Distância entre Eixos</i>	<i>Comprimento</i>	<i>Largura</i>	<i>Peso sem Carga</i>	<i>Cilindrada</i>	<i>Diâmetro Interno</i>	<i>Curso</i>
Executivo	4	6	10120	17680	6480	2765	164	331	319
Executivo	4	5	10580	19270	7140	2954	136	319	340
Econômico	4	4	9720	17200	6540	2340	108	362	264
Econômico	2	4	9450	16530	6380	1889	97	315	329
Econômico	2	4	9370	15000	6400	1837	79	291	307
<i>Taxa de Compressão</i>	<i>Potência</i>	<i>RPM Máximo</i>	<i>Consumo Urbano (MPG)</i>	<i>Consumo Rodoviário (MPG)</i>	<i>Decision</i>				
900	121	4250	21	28	4				
850	110	5500	19	25	3				
900	94	5200	26	32	4				
940	69	5200	31	37	3				
1010	60	5500	38	42	3				

Fonte: Schlimmer (1985)