



Pós-Graduação em Ciência da Computação

**SISTEMA AUTOMÁTICO PARA NEGOCIAÇÃO DE
AÇÕES USANDO TÉCNICA DE MINERAÇÃO DE DADOS
COM DETECÇÃO DE MUDANÇA DE CONCEITO**

Por

Victor Lorena de Farias Souza

Dissertação de Mestrado



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE

2015



Universidade Federal de Pernambuco
Centro de Informática
Pós-graduação em Ciência da Computação

Victor Lorena de Farias Souza

**SISTEMA AUTOMÁTICO PARA NEGOCIAÇÃO DE AÇÕES USANDO
TÉCNICA DE MINERAÇÃO DE DADOS COM DETECÇÃO DE
MUDANÇA DE CONCEITO**

*Trabalho apresentado ao Programa de Pós-graduação em
Ciência da Computação do Centro de Informática da Univer-
sidade Federal de Pernambuco como requisito parcial para
obtenção do grau de Mestre em Ciência da Computação.*

Orientador: Prof. Adriano Lorena Inácio de Oliveira

RECIFE
2015

Catálogo na fonte
Bibliotecária Jane Souto Maior, CRB4-571

S729s Souza, Victor Lorena de Farias

Sistema automático para negociação de ações usando técnica de mineração de dados com detecção de mudança de conceito / Victor Lorena de Farias Souza – Recife: O Autor, 2015.

105 f.: il., fig., tab.

Orientador: Adriano Lorena Inácio de Oliveira.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, 2015.

Inclui referências e apêndice.

1. Inteligência artificial. 2. Mineração de dados. 3. Sistemas de reconhecimento de padrões. I. Oliveira, Adriano Lorena Inácio de (orientador). II. Título.

006.3

CDD (23. ed.)

UFPE- MEI 2015-187

Dissertação de Mestrado apresentada por **Victor Lorena de Farias Souza** à Pós Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título “**SISTEMA AUTOMÁTICO PARA NEGOCIAÇÃO DE AÇÕES USANDO TÉCNICA DE MINERAÇÃO DE DADOS COM DETECÇÃO DE MUDANÇA DE CONCEITO**”, orientada pelo **Prof. Adriano Lorena Inácio de Oliveira** e aprovada pela Banca Examinadora formada pelos professores:

Prof. Aluizio Fausto Ribeiro Araújo
Centro de Informática / UFPE

Prof. Tiago Alessandro Espinola Ferreira
Departamento de Estatística e Informática / UFRPE

Prof. Adriano Lorena Inácio de Oliveira
Centro de Informática / UFPE

Visto e permitida a impressão.
Recife, 19 de outubro de 2015.

Profa. Edna Natividade da Silva Barros

Coordenadora da Pós-Graduação em Ciência da Computação do
Centro de Informática da Universidade Federal de Pernambuco.

Dedico este trabalho a todos os meus familiares, amigos e professores que contribuíram para sua realização.

Agradecimentos

Acima de tudo agradeço a Deus pela motivação e pela força necessárias ao longo dessa jornada. Agradeço, ainda, a todos aqueles que me apoiaram no desenvolvimento deste trabalho.

Um agradecimento especial aos meus pais Jorge Souza e Débora Farias, aos meus irmãos Jorge Lorena e Roberta Lorena e à minha namorada, Ludymila Oliveira, por sempre acreditarem em mim, me incentivarem e me apoiarem em todos os momentos da realização deste trabalho.

Agradeço ainda às seguintes pessoas:

- A todos da minha família;
- Ao Rodrigo Brasileiro, por toda a sua contribuição;
- Ao Professor Adriano Lorena, pelo apoio, confiança, oportunidade, críticas construtivas, orientações e excelentes conselhos fornecidos e nestes últimos anos;
- A todos os amigos que fiz no decorrer do curso, em especial, Marcos Falcão e Henrique Dantas;
- Aos professores do Centro de Informática pelos ensinamentos passados;
- Aos professores Aluizio Fausto Ribeiro Araújo e Tiago Alessandro Espinola Ferreira que aceitaram fazer parte da banca examinadora.

If you try to remember you will lose! Empty your mind. Be formless, shapeless, like water. Put water into a cup, it becomes the cup. Put water into a teapot, it becomes the teapot. Water can flow or creep or drip or crash.

Be water, my friend.

—BRUCE LEE

Resumo

Uma série temporal financeira representa as cotações dos preços das ações e apresenta comportamento similar a um fluxo de dados. Para a descoberta de padrões presentes em seus dados alguns trabalhos utilizam técnicas de mineração de dados que são fundamentadas na ideia de que os dados históricos guardam a memória essencial para prever a direção futura dos preços. Métodos tradicionais propostos na literatura consideram que o ambiente é estático, ou seja, que o mecanismo gerador da série financeira é o mesmo durante todo o intervalo de tempo de interesse. Porém, no caso de séries temporais financeiras, isso pode não ocorrer. Para resolver este problema, esta dissertação propõe a abordagem PAA-IDPSO-CD (Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Auto Adaptativa com detecção de mudança de conceito) para descoberta de padrões em séries temporais financeiras. A abordagem proposta objetiva lidar explicitamente com mudanças de conceito na série e descobrir os melhores padrões representativos dos dados das séries temporais que serão utilizados junto a uma estratégia de investimento formulada para automatizar as operações a serem feitas no mercado de ações. Isso possibilitará a redução das incertezas e dos riscos envolvidos nas compras e vendas de ações e auxiliará os investidores a maximizar o lucro nas suas operações feitas no mercado de ações. A fim de alcançar melhores resultados são propostas diferentes estruturas de partículas, utilizadas pelo IDPSO, junto a diferentes regras de decisão. Primeiramente, é utilizada uma estrutura básica para a partícula, em que se opera apenas na posição comprado no mercado financeiro. É proposto também uma segunda estrutura que é capaz de operar tanto na posição comprado quanto na posição vendido. Os experimentos do presente estudo comparam os resultados das versões do método proposto entre si e com os resultados obtidos pelas abordagens Buy and Hold (B&H) e SAX-GA (Aproximação por Valor Agregado Simbólico - Algoritmos Genéticos). Para isso, foram realizados Teste t Pareado com nível de confiança de 95% em vinte ações. O presente estudo conclui que o PAA-IDPSO-CD apresentou resultados estatisticamente melhores que o B&H e o SAX-GA para todas as vinte ações em que os testes foram executados ($p - valor < 0,05$). Além disso, a estratégia que opera nas posições comprado e vendido é melhor quando comparada àquela que opera apenas na posição comprado. No estudo comparativo em onze ações não houve diferença estatística e em outras sete a estratégia comprado e vendido obteve melhores resultados ($p - valor < 0,05$).

Palavras-chave: Descoberta de padrões. Reconhecimento de padrões. Mineração de dados. Representação PAA. Otimização por enxame de partículas. Detecção de mudança de conceito. Mercado de ações.

Abstract

Financial time series represents the prices of stock over time and presents similar behavior to a data stream. Mining techniques, which are based on the idea that the historical data retain the essential memory to predict the future direction, are used to make the patterns discovery in the time series data. In this context, this study proposes the PAA-IDPSO-CD (Piecewise Aggregate Approximation - Improved self-adaptive particle swarm optimization with Concept Drift) approach aiming to find patterns in financial time series. The purpose of this study is the discovery of the best representative patterns of the time series data to be used by a designed investment strategy to automate the operations to be made in the stock market. Thus, reducing the uncertainties and risks involved in buying and selling operations of stocks and help investors maximize the profit in their operations made in the stock market. In order to achieve best results, different particle structures, used by IDPSO, along with different decision rules are proposed. First, a basic structure to the particle is employed, it operates only in the Long position in the financial market. Subsequently, this structure develops to be able to operate both in Long and Short position. The experiments of this study comparing the results of the various versions of the proposed approach with each other and with those obtained by the Buy and Hold (B&H) and SAX-GA techniques, for that, were performed Paired t test with a confidence level of 95% in twenty stocks. The achieved results show that the PAA-IDPSO-CD outperforms B&H and SAX-GA for all twenty stocks in which the tests were performed ($p - value < 0.05$). Furthermore, the long-short operating strategy is better than that operating only in the long position. In the comparative study in eleven actions there was no statistical difference and in other seven long-short strategy outperforms ($p - value < 0.05$).

Keywords: Pattern discovery. Pattern recognition. Data mining. PAA representation. Particle swarm optimization. Concept drift. Stock market.

Lista de Figuras

2.1	Padrão triângulo simétrico.	22
2.2	Padrão triângulo ascendente.	24
2.3	Padrão triângulo descendente.	25
2.4	Padrão cabeça-e-ombros	26
2.5	Padrão cabeça-e-ombros invertido	27
2.6	Padrão topo duplo	28
2.7	Padrão fundo duplo	29
2.8	Representação PAA	33
2.9	PAA intervalos não interiores	34
2.10	Representação do cálculo MINDIST para PAA	35
2.11	Representação SAX	36
3.1	Evolução do modelo proposto	49
3.2	Estrutura da partícula operando apenas em posições comprado	51
3.3	Diagrama representando a regra de negócio operando apenas em posições comprado.	52
3.4	Estrutura da partícula operando nas posições comprado e vendido	53
3.5	Diagrama representando a regra de negócio para operar em posições comprado e vendido.	54
3.6	Cálculo da métrica de detecção de mudança de conceito proposta nas janelas de treinamento e retreinamento	59
3.7	Detecção de mudança de conceito no conjunto de teste	60
3.8	Particionamento do saldo inicial para redução de variância do modelo.	61
4.1	Convergência PSO com 20, 50 e 100 partículas.	66
4.2	Convergência Algoritmo Genético com 50 e 500 indivíduos.	66
4.3	Boxplot do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação AA	69
4.4	Boxplot do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação ABEV3	69
4.5	Boxplot do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação ITSA4	70
4.6	Boxplot do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação USIM5	70
4.7	Padrão PAA representativo da série temporal AA, em uma execução	72
4.8	Saldo do modelo ao longo do uso do sistema a partir do padrão PAA apresentado	73

4.9	Boxplots do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações AA, AAPL, ABEV3 e BBDC3	79
4.10	Boxplots do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações CRUZ3, ELET3, EMBR3 e IBM	80
4.11	Boxplots do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações ITSA4, KO, LIGT3 e MCD . . .	81
4.12	Boxplots do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações MMM, NKE, PETR3 e POMO4	82
4.13	Boxplots do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações SBUX, T, TIF e USIM5	83
4.14	Boxplot do PAA-IDPSO-CD operando em posições comprado-vendido, com e sem partição do saldo inicial, nas 50 execuções para a ação AA.	88
A.1	Série temporal das ações AA, AAPL, ABEV3 e BBDC3	99
A.2	Série temporal das ações CRUZ3, ELET3, EMBR3	100
A.3	Série temporal das ações IBM, ITSA4, KO	101
A.4	Série temporal das ações LIGT3, MCD, MMM	102
A.5	Série temporal das ações NKE, PETR3, POMO4	103
A.6	Série temporal das ações SBUX, T, TIF	104
A.7	Série temporal da ação USIM5	105

Lista de Tabelas

2.1	Pontos de Interrupção dividindo uma distribuição de Gauss em um número (3-5) regiões equiprováveis.	35
4.1	Comparativo dos resultados obtidos pelo PAA-PSO e SAX-GA nas 50 execuções.	68
4.2	Testes estatísticos para o comparativo dos resultados obtidos pelo PAA-PSO e SAX-GA nas 50 execuções, com nível de confiança de 95%.	71
4.3	Comparativo dos resultados obtidos pelo PAA-PSO e PAA-IDPSO, sem detecção de mudança de conceito, para as 50 execuções.	74
4.3	Continuação.	75
4.4	Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO e PAA-IDPSO, sem detecção de mudança de conceito, nas 50 execuções, com nível de confiança de 95%	75
4.5	Comparativo dos valores obtidos pelo PAA-PSO-CD E PAA-IDPSO-CD, para as 50 execuções.	77
4.5	Continuação.	78
4.6	Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO e PAA-IDPSO, na comparação dos modelos com e sem detecção de mudança de conceito, nas 50 execuções, com nível de confiança de 95%.	79
4.7	Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO-CD e PAA-IDPSO-CD, nas 50 execuções, com nível de confiança de 95%.	84
4.8	Comparativo dos valores obtidos pelo PAA-IDPSO-CD comprado-vendido com e sem partição de saldo inicial, para as 50 execuções.	87
4.9	Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-IDPSO operando em comprado-vendido com e sem partição de saldo inicial, nas 50 execuções, com nível de confiança de 95%.	88

Lista de Abreviaturas e Siglas

ABC	Colônia Artificial de Abelhas (ABC)
ANN	Rede Neural Artificial (Artificial Neural Network)
DDM	Método de Detecção de Mudança (Drift Detection Method)
DTW	Alinhamento Dinâmico no Tempo (Dynamic Time Warping)
EDDM	Método Detecção Precoce de Mudança (Early Drift Detection Method)
FT	Transformada de Fourier (Fourier Transform)
FFT	Transformada Rápida de Fourier (Fast Fourier Transform)
GA	Algoritmos Genéticos (Genetic Algorithm)
IDPSO	Otimização por Enxame de Partículas Auto Adaptativa (Improved self-adaptive particle swarm optimization)
KNN	Classificação dos K Vizinhos Mais Próximo (K-NN)
PAA	Aproximação por Valor Agregado de Segmento (Piecewise Aggregate Approximation)
PAA-IDPSO	Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Auto Adaptativa
PAA-IDPSO-CD	Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Auto Adaptativa com detecção de mudança de conceito
PAA-PSO	Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas
PSO	Otimização por Enxame de Partículas (Particle Swarm Optimization)
SAX	Aproximação por Valor Agregado Simbólico (Symbolic Aggregate Approximation)
SAX-GA	Aproximação por Valor Agregado Simbólico - Algoritmos Genéticos
SVD	Decomposição por Valor Singular(Singular Value Decomposition)
SVR	Regressão de Vetores de Suporte (Support Vector Regression)

Sumário

1	Introdução	15
1.1	Contexto e Motivação	17
1.2	Objetivos	18
1.3	Estrutura do Documento	18
2	Fundamentos	20
2.1	Introdução	20
2.2	Padrões Gráficos em Mercados Financeiros	20
2.3	Séries Temporais	29
2.3.1	Mineração de séries temporais	30
2.3.1.1	Estado da arte: Técnicas de representação de séries temporais aplicadas ao mercado financeiro	30
2.3.1.2	Aproximação por Valor Agregado de Segmento (PAA)	32
2.3.1.3	Aproximação por Valor Agregado Simbólico (SAX)	34
2.4	Mudança de conceito (<i>concept drift</i>)	36
2.4.1	Estado da arte: Técnicas de detecção de mudança de conceito	37
2.5	Otimização	39
2.5.1	Estado da Arte: otimização aplicada ao mercado financeiro	39
2.5.2	Algoritmo Genético	41
2.5.3	Otimização Por Enxame de Partículas	43
2.5.4	Otimização por Enxame de Partículas Auto Adaptativa	44
2.6	Considerações finais	46
3	Método Proposto	48
3.1	Introdução	48
3.2	PAA-IDPSO operando apenas em posição comprado	48
3.3	PAA-IDPSO operando nas posições comprado e vendido	52
3.4	PAA-IDPSO-CD	56
3.4.1	Métrica de detecção de mudança de conceito proposta	57
3.4.2	Método para redução da variância	60
3.5	Considerações finais	61
4	Experimentos	62
4.1	Introdução	62
4.2	Experimentos	63

4.2.1	Comparação PAA-PSO com SAX-GA ambas com regra de negócio operando apenas em posição comprado	67
4.2.2	Comparação PAA-IDPSO e PAA-PSO sem retreino	73
4.2.3	Comparação PAA-IDPSO-CD e PAA-PSO-CD com retreino baseado na detecção de mudança de conceito	76
4.2.4	Comparação PAA-IDPSO-CD comprado-vendido com e sem partição de saldo inicial	85
4.3	Considerações finais	89
5	Conclusão	90
5.1	Contribuições	91
5.2	Restrições da abordagem proposta	91
5.3	Trabalhos futuros	92
	Referências	93
	Apêndice	98
A	Séries temporais das ações utilizadas	99

1

Introdução

Uma série temporal é um conjunto de observações ordenadas sob algum parâmetro (BOX; JENKINS; REINSEL, 2011), por exemplo, as séries temporais financeiras seguem uma ordem cronológica (KANNAN et al., 2010). Estas representam as cotações dos preços das ações no mercado financeiro expressos em unidades de tempo, que podem ser diárias, semanais, mensais, anuais, dentre outros. Assim, uma série temporal financeira apresenta comportamento similar a um fluxo de dados, ou seja, como um conjunto de elementos de entrada que chegam de forma contínua e sequencial no tempo (GAMA, 2012).

Dentre as características marcantes da natureza dos dados das séries temporais financeiras estão a sua alta dimensionalidade (isto é, grande número de dados) e a necessidade de atualização contínua. Para o estudo desses dados e de suas características podem ser utilizadas técnicas de mineração de dados (FU, 2011). Devido ao aumento no tempo de execução e de espaço de armazenamento das operações a serem efetuadas, a alta dimensionalidade dos dados das séries temporais é um dos principais problemas a serem lidados na mineração desses dados. Entre as abordagens mais comuns para resolver esta dificuldade estão os Métodos de Representação, que permitem 1) a redução de dimensionalidade dos dados e 2) o cálculo de correlação e similaridade entre séries temporais, a partir da representação da série original em outro domínio com menos dimensões (FUAD; MARTEAU, 2013).

A mineração de dados pode ser descrita como “fazer um melhor uso dos dados” e é fundada na teoria de que os dados de histórico guardam a informação para estimar seu comportamento futuro. No cenário de mercado financeiro, a mineração de dados é utilizada na descoberta de padrões ocultos nos dados históricos das séries temporais, que podem representar seu comportamento, suas tendências, sua sazonalidade entre outras informações (ARGIDDI; APTE, 2014). Dessa forma, esses padrões podem ser utilizados para determinar pontos de compra e venda de ações e, conseqüentemente, auxiliarem os investidores em sua tomada de decisão (KANNAN et al., 2010).

No entanto, a dinamicidade, como característica inerente em séries temporais, pode implicar em mudanças na distribuição dos dados e em seus padrões, ao longo do tempo. Isso representa um grande desafio para os algoritmos tradicionais de aprendizagem, em que não há

retreino ao longo do tempo de uso dos modelos, já que os modelos aprendidos podem se tornar obsoletos (GAMA et al., 2014). Nesse contexto, podem ser aplicadas técnicas de detecção de mudança de conceito para detectar tais mudanças.

Em mineração de séries temporais, a Aproximação por Valor Agregado Simbólico (SAX) LIN et al. (2003) e a Aproximação por Valor Agregado de Segmento (PAA) KEOGH et al. (2001) são técnicas de representação e de redução da dimensionalidade e quantidade dos dados das séries temporais. Uma característica do SAX é ele ser um algoritmo baseado no PAA. No entanto, os valores representativos do SAX são discretos e do PAA são contínuos (LIN et al., 2007). Ambas as técnicas utilizam a métrica MINDIST (LIN et al., 2007), que representa uma distância euclidiana utilizada para o cálculo de proximidade entre suas representações. Essa métrica representa o limite inferior da distância Euclidiana e, por isso, permite executar os algoritmos de mineração de dados sobre a representação simbólica, de forma eficiente, e produzir resultados idênticos aos algoritmos que operam sobre os dados originais (LIN et al., 2003). O uso de ambas as técnicas pode acelerar o processo de extração de dados em séries temporais, mantendo a qualidade dos resultados de mineração. Como consequência de suas características, o PAA e o SAX vêm sendo utilizado em vários domínios de aplicação, entre eles, na área financeira (CANELAS; NEVES; HORTA, 2012).

Recentemente, com o objetivo de utilizar a representação SAX no mercado financeiro, CANELAS; NEVES; HORTA (2012) e CANELAS; NEVES; HORTA (2013) apresentaram modelos estáticos que combinam o SAX com o Algoritmo Genético (GA), resultando no algoritmo SAX-GA. Os autores utilizaram o seu modelo no contexto de descoberta de padrões em séries temporais financeiras, com o objetivo de identificar padrões representativos das séries temporais a serem utilizados junto à regra de negócio proposta em operações a serem feitas no mercado de ações. Um ponto fraco apresentado pelo SAX-GA é o fato de ele não ser dinâmico, ou seja, ocorre o treinamento e depois disso o sistema é usado sem alterações dos padrões descobertos e dos demais parâmetros do modelo. Na prática, isso pode ocasionar resultados ruins, uma vez que podem ocorrer mudanças no fenômeno gerador da série temporal ao longo do uso do modelo. Portanto, seria desejável usar modelos que detectem tais mudanças e se adaptem.

O presente estudo propõe a técnica PAA-IDPSO-CD (Aproximação por Valor Agregado de Segmento (PAA) - Otimização por Enxame de Partículas Auto Adaptativa (IDPSO) com Detecção de Mudança de Conceito (CD)). Este é um modelo dinâmico, que combina a técnica de representação e redução de dimensionalidade de dados PAA com o algoritmo Otimização por Enxame de Partículas Auto Adaptativa (IDPSO) (ZHANG; XIONG; ZHANG, 2013). O modelo proposto busca descobrir padrões que são utilizados junto à regra de negócio proposta, para a criação de estratégias de investimento que auxiliem os investidores a maximizar seus lucros nas suas operações em mercado de ações. Para isso, o modelo também utiliza uma técnica de detecção de mudança de conceito proposta, como gatilho para identificação de que o padrão utilizado está obsoleto e que o modelo precisa ser retreinado.

Para a detecção de mudança de conceito, o presente estudo utilizou a métrica *Drawdown*

(JOHANSEN, 2003). O *drawdown* é definido como a queda relativa no preço de um máximo local para o próximo mínimo local, ignorando as alterações de preços entre os dois valores extremos. No PAA-IDPSO-CD, o *drawdown* sobre o saldo obtido pode ser entendido como o erro do modelo, pois a queda no saldo ocorre quando o modelo não dispara uma ordem de compra ou de venda no melhor momento. Dessa forma, durante o treinamento calcula-se o *Máximo Drawdown* (maior erro apresentado pelo modelo), que será utilizado como limiar para a realização do retreino na fase de teste. Se durante a etapa de teste o saldo resultante do modelo cair a um valor maior que o *Máximo Drawdown* do treinamento, será ativado o retreino e serão calculados, a partir de uma nova otimização, novos padrões e demais parâmetros para o modelo além de um novo valor de *Máximo Drawdown*, a partir dos dados mais recentes da série temporal.

Até chegar a versão final, ou seja, o PAA-IDPSO-CD, foram desenvolvidos os métodos PAA-PSO (Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Padrão) e PAA-IDPSO (Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Auto Adaptativa), modelos que não apresentam o retreino, tendo como diferencial entre si o algoritmo de otimização utilizado. Por sua vez, o PAA-PSO-CD (Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Padrão com Detecção de Mudança de Conceito) e o PAA-IDPSO-CD são as respectivas versões dos modelos combinados com o retreino baseado na detecção de mudança de conceito.

Para validar a eficácia deste modelo, inicialmente os resultados obtidos pelo PAA-PSO são comparados em termos de lucratividade e tempo de execução com aqueles apresentados por SAX-GA. Posteriormente, o PAA-IDPSO é comparado com o PAA-PSO pelos mesmos critérios, ambos rodando em suas versões com e sem retreino em ambas as estratégias de investimento (operando apenas em posição comprado ou operando nas posições comprado e vendido) com o intuito de demonstrar o resultado final do modelo. Os resultados ainda são comparados com aqueles obtidos pela estratégia *Buy-and-Hold* que consiste em comprar as ações, mantendo-as em sua carteira de investimentos por um longo período de tempo. Esta estratégia é amplamente utilizada como uma referência de comparação nos estudos aplicados ao mercado financeiro (TEIXEIRA; OLIVEIRA, 2010), (BRASILEIRO et al., 2013), (CANELAS; NEVES; HORTA, 2012), (CANELAS; NEVES; HORTA, 2013).

1.1 Contexto e Motivação

A principal motivação para a abordagem proposta é a maximização do lucro nas operações feitas no mercado de ações através da redução das incertezas e dos riscos envolvidos nas compras e vendas de ações realizadas nesse mercado. Com o uso da abordagem proposta, as chances de lucro devem ser aumentadas à medida que o modelo seja capaz de definir bons momentos para a realização de operações de compra e venda de ações e de oferecer informações para auxiliarem os investidores em sua tomada de decisão. Assim, a criação de soluções que sirvam como um sistema de suporte de decisão, de forma confiável, para as operações realizadas

nesse setor da economia é bem recebida pelos investidores da comunidade financeira.

Além disso, o uso de detecção de mudança de conceito e, posterior, retreino é de fundamental importância para esse mercado. Uma vez que, ele trabalha com séries temporais e estas são fluxos de dados dinâmicos que podem mudar o seu comportamento com o passar do tempo. Nesse cenário, está envolvido o desafio de criar um sistema que possa lidar e se adaptar, de forma rápida, às mudanças de comportamento das ações.

1.2 Objetivos

O objetivo deste trabalho é o desenvolvimento de um método de mineração de dados utilizado na descoberta de padrões que possam ser utilizados junto a regras de decisão para servir como um sistema de suporte de decisão das operações a serem feitas no mercado de ações e que lide explicitamente com a mudança de conceitos nas séries temporais. A ideia é a geração lucro financeiro nas operações feitas pelo método, no mercado de ações, com risco controlado. Esse controle é realizado a partir da detecção de mudança de conceito baseado no *Drawdown*, que é utilizado na prática como métrica de risco do mercado financeiro.

Para isso, o modelo proposto utiliza a técnica de representação e redução de dimensionalidade de dados PAA, que, quando combinado ao algoritmo de otimização IDPSO e à detecção de mudança de conceito, efetua a busca pelo melhor padrão representativo das séries temporais financeiras. Esses padrões são utilizados junto a regras de decisões que têm seus parâmetros redefinidos ao longo da série, através do retreino decorrente da detecção de mudança de conceito.

O modelo proposto foi testado em ações tanto do mercado norte americano (índice S&P 500) quanto do mercado brasileiro (BM&FBOVESPA), em condições reais de mercado. Os resultados são comparados entre os obtidos pelas diversas versões apresentadas do modelo proposto até se chegar a sua versão final e, também, pelas estratégias *Buy-and-Hold* e *SAX-GA*. A ideia é mostrar que a versão final da abordagem proposta é superior às demais em relação ao saldo final do período (lucro ou perda).

1.3 Estrutura do Documento

Esta dissertação está estruturada em 6 capítulos. O primeiro é este introdutório e os demais estão descritos na sequência.

No capítulo 2 são contextualizados os assuntos padrões gráficos utilizados em mercados financeiros, séries temporais, técnicas de representação dos dados utilizadas em mineração de séries temporais, com foco nas abordagens PAA e SAX, detecção de mudança de conceito e algoritmos de otimização, dando ênfase aos algoritmos GA, PSO e IDPSO.

No capítulo 3 é apresentado o modelo proposto pelo presente estudo, o PAA-IDPSO-CD. São discutidas e justificadas as técnicas utilizadas para a realização da mineração das séries temporais, isto é, a descoberta de padrões, da otimização e da detecção da mudança de conceito.

Além disso, são descritos outros métodos desenvolvidos ao longo do trabalho e sua evolução até a versão final, ou seja, o PAA-IDPSO-CD.

O capítulo 4 apresenta os experimentos. Esse capítulo apresenta os resultados obtidos pelas diversas versões apresentadas pelo modelo proposto até se chegar a sua versão final. Os resultados dos modelos são comparados entre si e com os obtidos pelas estratégias *Buy-and-Hold* e SAX-GA.

No capítulo 5 são apresentadas as conclusões e os trabalhos futuros. Também são feitas observações sobre as contribuições feitas e as limitações dos métodos propostos.

2

Fundamentos

2.1 Introdução

Neste capítulo são contextualizadas as áreas de pesquisa 1) padrões gráficos utilizados em mercados financeiros, 2) séries temporais, com foco nas técnicas de representação utilizadas em mineração de dados de séries temporais, 3) detecção de mudança de conceito (*concept drift*), e 4) otimização, que são as abordagens computacionais por trás do método proposto.

Nos padrões gráficos usados em mercados financeiros, são destacados os padrões de continuação e de reversão de tendência e os cenários em que são utilizados. Além disso, são apresentadas algumas dificuldades e vantagens para o seu uso.

Nas técnicas de representação utilizadas em mineração de dados aplicada a séries temporais serão apresentadas um breve estado da arte abordando o seu uso em aplicações do mercado financeiro. Será dada ênfase nas técnicas PAA e SAX, que são respectivamente a técnica utilizada pelo presente estudo e a abordagem de comparação.

Por sua vez, na detecção de mudança de conceito serão apresentados seus os conceitos além de um estado da arte das técnicas mais utilizadas atualmente.

Na sequência, são apresentados o estado da arte sobre o uso de algoritmos de otimização em modelos utilizados no Mercado de Ações em suas várias formas. Além disso, ainda são descritos o GA, o PSO e o IDPSO, este último como o algoritmo utilizado pelo modelo proposto, sendo as demais utilizadas para comparação. Por fim, são apresentados as considerações finais.

2.2 Padrões Gráficos em Mercados Financeiros

Os padrões gráficos (chart patterns) de mercados financeiros são utilizados, principalmente, pelos investidores analistas técnicos. A análise técnica estuda, a oferta e a demanda de uma ação no mercado em uma tentativa de determinar em que direção, ou tendência, sua cotação irá seguir no futuro e se baseia em três premissas: (1) Preço reflete os demais fatores presentes no mercado; (2) Preços das ações se movem em tendências; (3) O comportamento histórico dos preços tende a se repetir (VANSTONE; FINNIE, 2009).

1. Preço reflete os demais fatores presentes no mercado. A análise técnica assume que, em um dado momento, o preço de uma ação já reflete os demais fatores que pode afetar a empresa. Os analistas técnicos acreditam que os fundamentos da empresa, juntamente com fatores econômicos mais amplos e a psicologia do mercado, são refletidos nos preços da ação, eliminando a necessidade de se considerar, esses fatores separadamente.

2. Preços das ações se movem em tendências. Para a análise técnica, os movimentos de preços seguem tendências. Isto significa que, depois de uma tendência ter sido estabelecida, é mais provável que o movimento do seu preço futuro se mantenha no mesmo sentido.

3. O comportamento histórico dos preços tende a se repetir. A análise técnica ainda assume que o comportamento histórico dos preços das ações tende a se repetir. A natureza repetitiva é atribuída à psicologia do mercado, isto é, os participantes do mercado tendem a proporcionar uma reação consistente para estímulos de mercado semelhantes ao longo do tempo. Dessa forma, a análise técnica usa padrões gráficos para analisar o comportamento dos movimentos do mercado e compreender as suas tendências.

Para muitos operadores do mercado de ações analistas técnicos, o gráfico de uma ação é o ponto de partida para as análises do seu preço futuro. E por um bom motivo: gráficos podem fornecer uma série de informações em um pequeno espaço de tempo (MURPHY, 1999). Ao observar o gráfico representante da cotação das ações de uma empresa, os usuários da informação, conhecido como *chart reader*, podem dizer se o valor das ações de uma empresa tem crescido ou caído ao longo do período. Com isso, pode ver o quão bem ou mal está a situação das empresas diante do mercado financeiro, podem, ainda, determinar a volatilidade das ações da empresa. Uma empresa cujos movimentos de subida e de descida nos preços são muito irregulares é claramente mais volátil do que uma empresa cujas ações se movem de forma relativamente suave ao longo do tempo (MURPHY, 1999).

Padrões gráficos podem sinalizar as tendências do mercado para os investidores, ou seja, que o preço de uma ação é suscetível de continuar a se mover na direção em que vinha se movendo ou revertê-la. Com isso, existem dois tipos de padrões na análise técnica: reversão e continuação. Um padrão de reversão sinaliza que a tendência vinha ocorrendo irá inverter, ao fim do padrão. Por sua vez, um padrão de continuação indica que a tendência anterior continuará a ocorrer (BULKOWSKI, 2005).

Antes de trabalhar com padrões gráficos é importante entender os conceitos de linhas de tendência que são linhas traçadas no gráfico para sinalizar um nível de suporte ou de resistência para o preço da ação. Linha de suporte representa um nível que, historicamente, os preços têm dificuldade de atingir patamares abaixo dela. Por outro lado, uma linha de resistência ilustra o nível em que os preços têm dificuldade de atingir patamares acima. Ambas as linhas podem ser níveis de preços constantes ou variar na direção da tendência ao longo do tempo (FISCHER; FISCHER, 2003).

A dificuldade na identificação de padrões e de como operá-los é que o uso do gráfico não é uma ciência exata. Na verdade, ele é, muitas vezes, visto como mais uma arte do que uma

ciência. Por conta disso, mesmo havendo uma idéia geral dos componentes e do comportamento do mercado para cada padrão gráfico, o movimento de preços pode não corresponder necessariamente ao padrão sugerido pelo gráfico. Isso não deve desencorajar os potenciais utilizadores de gráficos, uma vez que os conceitos básicos de criação de gráficos já são compreendidos e sua qualidade pode ser reforçada através do uso de indicadores de preço e de volume (BULKOWSKI, 2005). Além disso, com o aumento do poder computacional, sistemas computacionais podem ser utilizados com o objetivo de descobrir padrões nas séries temporais, apresentando como vantagens o fato de o sistema ser capaz de descobrir padrões específicos baseados no comportamento de cada ação (FU, 2011).

- **Padrão de continuação de tendência**

Um padrão de continuação de tendência sugere que uma tendência que vinha ocorrendo, e venha a apresentar um desvio temporário no comportamento, eventualmente, irá continuar a sua tendência anterior.

Entre os padrões mais conhecidos de continuação de tendência estão os padrões triângulos. Estes padrões gráficos usam a convergência das duas linhas de tendência, sejam elas planas, ascendentes ou descendentes, e o preço da ação em movimento entre as duas linhas de tendência. Existem três tipos de triângulos, que variam em construção e significado: o triângulo simétrico, o triângulo descendente e no triângulo ascendente (MURPHY, 1999).

- Padrão Triângulo Simétrico



Figura 2.1: Padrão triângulo simétrico.

O triângulo simétrico é considerado como um padrão de continuação que sinaliza um período de consolidação de uma tendência seguida por uma retomada da tendência anterior. Ele

é formado pela convergência de uma linha de resistência decrescente e uma linha de suporte crescente. As duas linhas de tendência para a formação desse triângulo devem convergir num ponto, conhecido como o ápice. O preço da ação apresentará volatilidade dentro dessas linhas, em direção ao ápice, e tipicamente sairá na direção da tendência anterior (FISCHER; FISCHER, 2003).

Se precedida por uma tendência descendente, o foco deve ser em uma ruptura abaixo da linha de suporte crescente. Se precedida de uma tendência de alta, deve ser observada uma quebra acima da linha de resistência decrescente. No entanto, esse padrão nem sempre leva a uma continuação da tendência anterior, uma ruptura na direção oposta da tendência anterior deve sinalizar a formação de uma nova tendência (FISCHER; FISCHER, 2003). A figura 2.1 representa um triângulo simétrico que é precedido por uma tendência de subida .

Como observado na figura 2.1 a primeira parte deste padrão é a criação de uma alta na tendência ascendente, que é seguida por uma queda nos preços. O preço, em seguida, move-se para uma outra elevação que é menor do que o primeiro pico, seguida por outra queda, menor que a primeira, resultando em um vale em posição mais elevada que o precedente. Neste ponto, as linhas de tendência pode ser tirada, criando seu ápice. O preço vai continuar a mover-se entre estas linhas até o ponto de ruptura (BULKOWSKI, 2005).

Assim, no caso de tendência ascendente, a cotação-alvo (T) é geralmente ajustada para ser igual ao preço de ruptura (R) mais a altura vertical do triângulo (H). Já na tendência descendente, a cotação-alvo é geralmente ajustada para ser igual ao preço de ruptura menos a altura vertical do triângulo (H). O cálculo da cotação alvo ocorre através das equações 2.1 e 2.2, (BULKOWSKI, 2005).

$$\text{Em caso de tendência ascendente: } T = R + H \quad (\text{Eq.2.1})$$

$$\text{Em caso de tendência descendente: } T = R - H \quad (\text{Eq.2.2})$$

- Padrão Triângulo Ascendente

O triângulo ascendente é um padrão de continuação de tendência de subida, sendo, normalmente, precedido por uma tendência ascendente. Isto dá uma indicação de que o preço da ação tende a ser maior após o fim do padrão. O padrão é formado por duas linhas de tendência: uma linha de tendência plana, sendo uma linha de resistência, e uma linha de suporte ascendente. O preço da ação se move entre estas linhas de tendência até finalmente rompê-las no sentido ascendente. A figura 2.2 é um exemplo de um triângulo ascendente (FISCHER; FISCHER, 2003).

Como pode ser observado na figura 2.2, o preço se move para um pico que enfrenta resistência levando a uma queda. Isto segue um outro movimento de subida, que testa o nível anterior de resistência. Após não conseguir mover-se após esse nível de resistência, o preço da ação novamente cai - mas para um vale em posição superior ao anterior. Isto continua até que



Figura 2.2: Padrão triângulo ascendente.

o preço consiga romper e alcance valores acima do nível de resistência ou o padrão de falha (BULKOWSKI, 2005).

O triângulo ascendente é utilizado da seguinte forma: operadores do mercado compram (vão operar na posição comprado) da ação, geralmente em alto volume, quando o seu preço consegue atravessar a linha de resistência e, com isso, aumentam agressivamente o preço da ação. Assim, a cotação-alvo (T) é geralmente ajustada para ser igual ao preço de ruptura (R) somado à altura vertical do triângulo (H). O cálculo da cotação alvo ocorre através da equação 2.3 (BULKOWSKI, 2005):

$$T = R + H \quad (Eq.2.3)$$

- Padrão Triângulo Descendente

O triângulo descendente é um padrão de continuação de tendência de descida, sendo, normalmente, precedido por uma tendência descendente. Isto sugere que o preço tende a ser menor após a conclusão do padrão. O triângulo descendente é construído com uma linha de suporte plana e uma linha de resistência com inclinação para baixo. O preço da ação se move entre estas linhas de tendência até finalmente rompê-las no sentido descendente. A figura 2.3 é um exemplo desse padrão (FISCHER; FISCHER, 2003).

Como pode ser observado na figura 2.3, a primeira parte deste padrão é a queda no preço por uma alta nos preços. O próximo passo é o segundo teste do nível de suporte anterior, que por sua vez envia o preço da ação para cima - mas para uma posição inferior do que o pico anterior. Este processo é repetido até que o preço é incapaz de manter o nível de suporte, rompendo-o e atingindo patamares inferiores, retomando a tendência descendente (BULKOWSKI, 2005).



Figura 2.3: Padrão triângulo descendente.

O triângulo descendente é utilizado da seguinte forma: operadores do mercado vão operar na posição vendido da ação quando o seu preço atinge patamares menores que a linha de suporte e, com isso, empurram o preço do ativo para baixo. Esse padrão dá aos operadores a oportunidade de fazer lucros substanciais no curto prazo. A cotação-alvo (T) é geralmente fixada para ser igual ao preço de ruptura (R) menos a altura vertical (H) entre as duas linhas de tendência (BULKOWSKI, 2005). A cotação-alvo pode ser calculada através da equação 2.4.

$$T = R - H \quad (Eq.2.4)$$

- **Padrão de reversão de tendência**

A reversão é uma mudança positiva ou negativa contra a tendência que vinha ocorrendo no preço da ação. Isto é, caso os preços de uma ação viessem seguindo uma tendência ascendente, após o fim do padrão de reversão, seus preços passariam a seguir uma tendência descendente, e vice-versa. Os operadores do mercado devem prestar atenção para esses padrões, pois podem indicar a necessidade de mudança na estratégia de negociação. Por exemplo, se um operador vinha operando em em posição comprado numa ação e percebe um padrão de reversão, ele deve considerar o fechamento de sua posição atual, a fim de assumir uma posição vendido para, com isso, lucrar sobre o movimento descendente potencial do preço da ação (MURPHY, 1999).

Entre os padrões de reversão mais conhecidos estão o cabeça-e-ombros (Head-and-Shoulders), cabeça-e-ombros invertido, topo duplo e fundo duplo.

- Padrão Cabeça-e-Ombros (Head and Shoulders)

O padrão de cabeça-e-ombros é um dos mais populares e confiáveis padrões gráficos em análise técnica. E como se poderia imaginar a partir do nome, o padrão se parece com uma cabeça com dois ombros.

Cabeça-e-ombros é um padrão de reversão que, quando formado, sinaliza a provável mudança da tendência que vinha sendo seguida pelo preço da ação. Esse padrão pode ser visto em duas versões, na primeira, o **padrão cabeça-e-ombros** sinaliza que o preço de uma ação tende a cair quando atingir o fim do padrão, e é geralmente formado no auge de uma tendência ascendente. A segunda versão, o **cabeça-e-ombros invertido**, sinaliza que o preço de uma ação tende a subir e, geralmente, se forma durante uma tendência de queda.

Ambos os padrões cabeça-e-ombros têm uma construção semelhante, em que há quatro partes principais: dois ombros, uma cabeça e um decote. Os padrões são confirmados quando a linha de decote é rompida, após a formação do segundo ombro.

O **padrão cabeça e os ombros** são conjuntos de picos e vales, sendo o decote um nível de suporte ou resistência. O padrão cabeça-e-ombros é baseado na análise da Teoria de Dow, ou seja, uma tendência ascendente, por exemplo, é vista como um período de sucessivos picos e vales em ascensão. Uma tendência descendente é entendida em situação contrária à anterior. O padrão de cabeça-e-ombros ilustra um enfraquecimento em uma tendência onde existe uma deterioração dos seus picos e vales (FISCHER; FISCHER, 2003).



Figura 2.4: Padrão cabeça-e-ombros.

Como pode ser observado na Figura 2.4, o padrão cabeça-e-ombros sinaliza para os operadores de mercado que o preço de uma ação irá fazer um movimento descendente, após o rompimento do nível de decote. Como este padrão é formado principalmente por picos de tendências ascendentes, ele é considerado um padrão de reversão de tendência, ou seja, o preço da ação tende a cair após o fim do padrão. A cotação-alvo (T) é igual ao preço no nível do decote (N) menos a altura vertical (H) entre o nível de decote e a cabeça (BULKOWSKI, 2005). A cotação-alvo pode ser calculada através da equação 2.5.

$$T = N - H \quad (Eq.2.5)$$

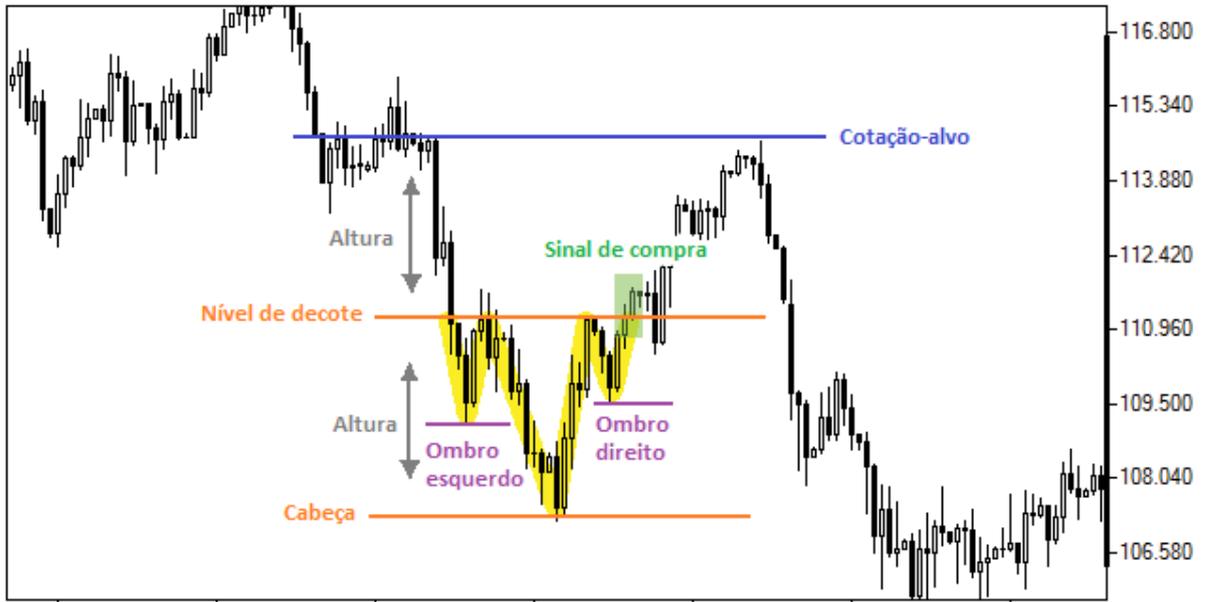


Figura 2.5: Padrão cabeça-e-ombros invertido.

Por sua vez, o **padrão cabeça-e-ombros invertido** é o oposto do mostrado anteriormente. Como pode ser visto na figura 2.5, o padrão cabeça-e-ombros invertido aponta que o preço das ações irá fazer um movimento ascendente após o rompimento do nível de decote. Por ocorrer ao fim de uma tendência descendente, o padrão é formado principalmente por vales, como ele é um padrão de reversão de tendência, o preço da ação tende a subir após o fim do padrão. A cotação-alvo (T) é igual ao preço no nível do decote (N) somado à altura vertical (H) entre o nível de decote e a cabeça (BULKOWSKI, 2005). A cotação-alvo pode ser calculada através da equação 2.6.

$$T = N + H \quad (Eq.2.6)$$

Em ambos os gráficos podem ser vistos os pontos fundamentais do padrão, seja ele invertido ou não, começando com a formação do ombro esquerdo; seguido pela a formação da cabeça, que é o segundo passo; o terceiro passo é a formação do ombro direito; por fim o padrão é completado quando há o rompimento do nível de decote.

- Padrão Topo Duplo e Fundo Duplo

O topo duplo e o fundo duplo são outro par de conhecidos padrões gráficos cujos nomes não deixa muito para a imaginação. Estes dois padrões de reversão ilustram a tentativa de uma ação continuar uma tendência existente. Após várias tentativas frustradas para se mover de forma a continuar a tendência anterior, esta é, então, invertida e uma nova começa. Estes padrões gráficos, muitas vezes, lembram o que se parece com um “W” (Fundo Duplo) ou um “M” (Topo Duplo).



Figura 2.6: Padrão topo duplo.

O **padrão topo duplo** é encontrada nos picos de uma tendência ascendente e é um sinal claro de que a tendência de alta anterior está enfraquecendo e que os operadores do mercado estão perdendo o interesse. Após a conclusão deste padrão, é provável que a tendência seja revertida e o preço da ação deverá cair (FISCHER; FISCHER, 2003).

Como pode ser visto na figura 2.6 a primeira etapa desse padrão é a criação de uma nova alta durante a tendência ascendente, que, após um pico, enfrenta resistência e cai a um nível de suporte. A próxima etapa deste padrão é o preço começar a se mover de volta para o nível de resistência encontrada anteriormente, que por sua vez volta para o nível de suporte. O padrão é completado quando o preço da ação rompe o nível de suporte, marcando, assim, o início de uma tendência descendente. A cotação-alvo (T) é igual ao preço no rompimento do nível de suporte (S) menos a altura vertical (H) entre o nível de suporte e o de resistência (BULKOWSKI, 2005). A cotação-alvo pode ser calculada através da equação 2.7.

$$T = S - H \quad (Eq.2.7)$$

Por sua vez, o **padrão fundo duplo** sinaliza uma reversão da tendência descendente para uma ascendente, ou seja, apresenta comportamento oposto ao padrão topo duplo.

Como observado na figura 2.7, o fundo duplo é formado quando uma tendência descendente define uma nova baixa no movimento dos preços. Esse movimento vai encontrar uma linha de suporte que impede uma queda maior no preço da ação. Ao encontrar a linha de suporte, a ação vai tender a uma nova alta, que atinge o seu ponto de resistência da ação. A próxima etapa deste padrão é outra queda no preço das ações. Estes dois testes de suporte formam os dois fundos no padrão gráfico. Mas, novamente, o preço ação tende a subir. O padrão é confirmado quando o preço se rompe o nível de resistência, marcando, assim, o início de uma tendência ascendente. A cotação-alvo (T) é igual ao preço no rompimento do nível de resistência (R)

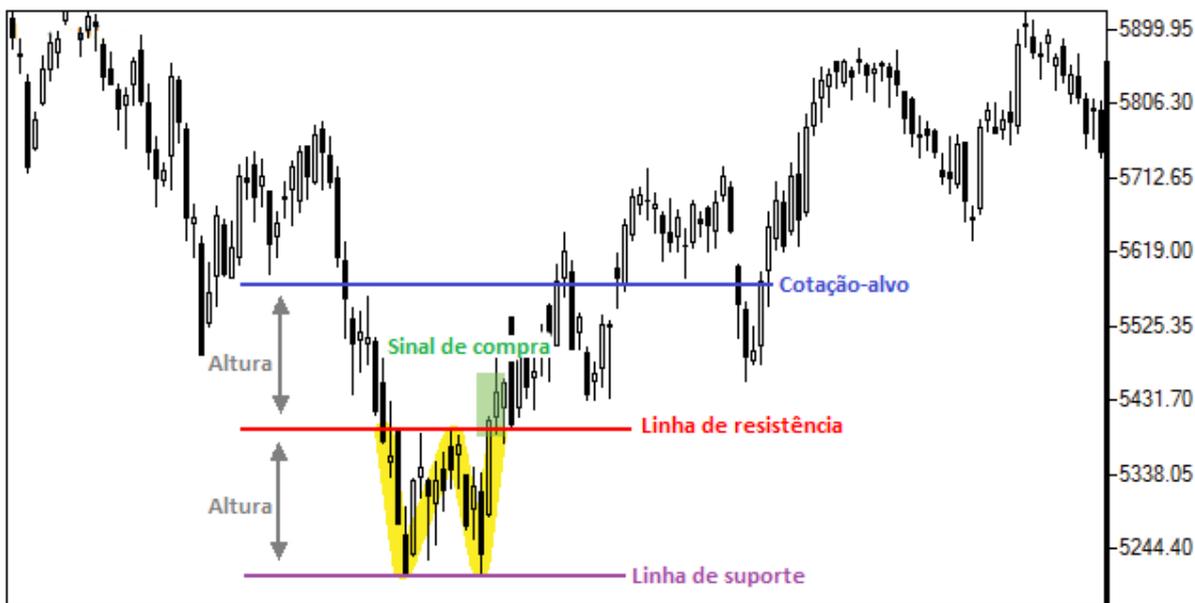


Figura 2.7: Padrão fundo duplo.

somado à altura vertical (H) entre o nível de suporte e o de resistência (BULKOWSKI, 2005). A cotação-alvo pode ser calculada através da equação 2.8.

$$T = R + H \quad (Eq.2.8)$$

Como após o padrão topo duplo é provável que ocorra uma tendência de queda, os operadores do mercado financeiro podem gerar lucro a partir de operações na posição vendido. Já no padrão fundo duplo, os lucros podem ser obtidos na posição comprado.

2.3 Séries Temporais

Uma série temporal é um conjunto de observações ordenadas sob algum parâmetro (BOX; JENKINS; REINSEL, 2011). Seus dados são utilizados para análise e compreensão do passado que podem ser utilizados como conhecimento para a previsão de seu comportamento futuro, permitindo que os usuários da informação possam tomar decisões devidamente informados. A análise de séries temporais quantifica suas principais características podendo identificar algum padrão nos seus dados ou, então, sua variação aleatória. Estas razões, combinadas com o maior poder de computação, fizeram os métodos de estudo das séries temporais serem amplamente utilizados pelos diversos setores da economia como, por exemplo, governo, indústria e comércio (COWPERTWAIT; METCALFE, 2009).

Como exemplo, as séries temporais financeiras apresentam as cotações dos preços das ações no mercado financeiro fixados em unidades de tempo, que podem ser, dentre outras, diárias, semanais, mensais ou anuais. Entre as principais características dos dados das séries temporais estão a sua alta dimensionalidade (grande número de dados) e necessidade de atualização

contínua. O uso de ferramentas de mineração de dados é uma das formas mais utilizadas para o estudo dessas informações (FU, 2011).

No contexto de mineração de dados de séries temporais, a tarefa mais comum é descobrir padrões a partir dos dados de séries temporais (FU, 2011). Por apresentarem um comportamento dinâmico, isto é, chegada de dados de forma contínua e sequencialmente no tempo (GAMA, 2012), as séries temporais podem apresentar mudanças na distribuição de seus dados e em seus padrões, ao longo do tempo. Essa característica representa um grande desafio para os algoritmos tradicionais de aprendizagem, em que não há retreino ao longo do tempo, já que os modelos aprendidos podem se tornar obsoletos (GAMA et al., 2014).

2.3.1 Mineração de séries temporais

A mineração de dados pode ser descrita como “fazer um melhor uso dos dados” e é fundada na teoria de que os dados de histórico guardam a informação para estimar seu comportamento futuro. No cenário de mercado financeiro, a mineração de dados é utilizada na descoberta de padrões ocultos nos dados históricos das séries temporais, que podem representar seu comportamento, suas tendências, sua sazonalidade entre outras informações (ARGIDDI; APTE, 2014). Dessa forma, esses padrões podem ser utilizados para determinar pontos de compra e venda de ações e, conseqüentemente, auxiliarem os investidores em sua tomada de decisão (KANNAN et al., 2010).

O principal objetivo da mineração de dados de séries temporais é descobrir informações ou conhecimentos ocultos a partir dos dados históricos das séries temporais, sejam elas em seu formato original ou transformadas. Como esses dados apresentam entre suas características a sua alta dimensionalidade e operar sobre eles seria muito custoso computacionalmente, entre os métodos mais comuns para lidar com esse problema estão as abordagens que realizam a representação desses em outros domínios que apresentem menos dimensões tornou-se um problema fundamental nessa área. Para que, nessas representações, possam ser calculadas as similaridade entre dados transformados como se fossem dados originais a um custo computacional menor (FU, 2011). Essa similaridade pode, então, ser utilizadas para a realização de outras tarefas, como, por exemplo, descoberta de padrões, classificação, agrupamento (FU, 2011).

Assim, nesta subseção, são apresentados um estado da arte sobre a utilização das técnicas de representação das séries temporais financeiras. Será dada uma maior ênfase na descrição das técnicas PAA no contexto de descoberta de padrões.

2.3.1.1 Estado da arte: Técnicas de representação de séries temporais aplicadas ao mercado financeiro

Dentre as técnicas utilizadas para a redução de dimensionalidade estão os seguintes: Transformada de Fourier (FT), Alinhamento Dinâmico no Tempo (DTW), Decomposição por Valor Singular (SVD), Aproximação por Valor Agregado de Segmento (PAA) e Aproximação

por Valor Agregado Simbólico (SAX) (SUN et al., 2014). Um ponto negativo do uso dessas técnicas é que decorrente da redução da dimensionalidade pode haver perda de informações úteis para a análise das séries temporais, mas seu uso combinado a outras abordagens podem reduzir essas perdas ou os seus efeitos (SUN et al., 2014).

DTW é uma técnica bem estabelecida para encontrar um alinhamento ótimo entre duas de séries temporais sob certas restrições. As sequências são geralmente transformadas em forma não lineares para coincidir uma com a outra. Essa abordagem tem sido usada com sucesso, entre outras áreas, para comparar diferentes sequências ou padrões e para recuperação de informações (THONGMEE et al., 2014).

THONGMEE et al. (2014) propõem o método Relacionamento Forte por Blocos (Blockwise Strong Relationship - BSR) que calcula a relação entre qualquer par de ações do mercado com base em apenas seus preços. Para isso o método utiliza a transformação dos dados a partir da técnica SAX e a medida de distância usando o DTW.

TSINASLANIDIS; KUGIUMTZIS (2014) propõem um algoritmo de previsão do preço de ações baseado na combinação de duas ferramentas de mineração de dados, a técnica Pontos Importantes Perceptualmente (Perceptually Important Points - PIP) usada para segmentar dinamicamente séries temporais financeiras em janelas menores e DTW usado para encontrar janelas de tempo similares. Em seguida, as previsões são feitas a partir dos mapeamentos entre as janelas mais semelhantes.

BAGHERI; PEYHANI; AKBARI (2014) apresentam um modelo híbrido que combina DTW com a abordagem Transformada Wavelet (WT) para extração automática de padrões das séries temporais financeiras, especialmente para o Mercado de Câmbio (FOREX). Os resultados indicam que o método apresentado é muito útil e eficaz para a previsão do preço e para a mineração de padrões das séries financeira.

De acordo com Fourier, qualquer função pode ser representada como uma soma de uma série de funções seno e cosseno. Em séries temporais, a Transformada de Fourier (FT) é utilizada para a análise espectral da série, ou seja, o estudo do comportamento da série em termos de uma composição média de frequência. Em termos mais gerais, a análise espectral pode ser usada para detectar sinais periódicos (COWPERTWAIT; METCALFE, 2009).

CHEN; CHEN (2014) propõem um novo método que integra técnica de discretização de entropia combinado com a técnica Transformada Rápida de Fourier (FFT) para desenvolver um novo modelo fuzzy de previsão de séries temporais. O modelo proposto é implementado usando o método de bootstrapping incrementalmente atualizado para aprimorar sua capacidade de previsão. Os resultados mostram que mecanismo de aprendizagem incremental do modelo proposto permite que ele efetivamente lidar com grandes conjuntos de dados financeiros de forma online.

ZHANG; XIONG; ZHANG (2013) se concentram em um modelo estatístico para precificação de ações baseado em um modelo de difusão salto exponencial duplo com volatilidade estocástica e taxa de juros estocástica. O modelo utiliza a Transformada Rápida de Fourier para

obter soluções numéricas para os preços das ações. Os experimentos mostram que para o modelo proposto, a FFT é rápida e eficiente.

HUANG; ZHU; RUAN (2014) propõem um modelo baseado na abordagem da Transformada Rápida de Fourier para a avaliação de opções quando as ações seguem o processo salto exponencial duplo com volatilidade estocástica e taxa de juros estocástica. Este modelo captura dois termos estruturais dos preços das ações: a volatilidade implícita do mercado e o comportamento salto. Os resultados numéricos dos experimentos mostraram que a abordagem FFT é uma abordagem rápida e que apresenta bons resultados.

Decomposição por Valor Singular (SVD) (BANERJEE; PAL, 2014) é uma técnica amplamente utilizada para decompor uma matriz em várias matrizes de componentes, expondo muitas das propriedades úteis e interessantes da matriz original. Assim, esta técnica é um método para redução de dados e pode ser entendido a partir de três pontos mutuamente compatíveis. Por um lado, pode ser visto como um método para transformar variáveis correlacionadas em um conjunto de variáveis não correlacionadas que apresentem melhor as várias relações entre os itens de dados originais. Ao mesmo tempo, SVD é um método para identificar e ordenar as dimensões ao longo de pontos de dados que apresentam a maior variação. Isto está relacionado com a terceira forma de ver o SVD, através dele é possível encontrar a melhor aproximação dos pontos de dados originais usando menos dimensões (BANERJEE; PAL, 2014).

No contexto de séries financeiras, CARAIANI (2014) propõe uma abordagem baseada em correlação para analisar os dados financeiros do mercado de ações dos Estados Unidos. Esse modelo utiliza cálculo da entropia com base na SVD da matriz de correlação para os componentes do índice Dow Jones.

Por sua vez, Canelas et al. propuseram dois modelos (CANELAS; NEVES; HORTA (2012); CANELAS; NEVES; HORTA (2013)) para automação de operações no mercado financeiro utilizando a técnica SAX. Em ambos os modelos, o SAX é combinado com um kernel de otimização baseada em algoritmos genéticos (GA) para identificar os padrões mais relevantes das séries temporais financeiras que são utilizados junto a regras de investimento para a automação das operações.

2.3.1.2 Aproximação por Valor Agregado de Segmento (PAA)

O algoritmo de Aproximação por Valor Agregado de Segmento (Piecewise Aggregate Approximation - PAA) foi proposto por KEOGH et al. (2001), sendo uma técnica de redução de dimensionalidade dos dados (numero de dados). Neste método, cada sequência de dados das séries temporais é dividida em k segmentos com igual comprimento e o valor médio de cada segmento é, então, utilizado como seu valor representativo. Ao final, a representação PAA da série temporal vai ser composta por um vetor k -dimensional das médias de seus segmentos. A Figura 2.8 mostra a representação PAA de uma série temporal Q , adaptado a partir de (LKHAGVA; SUZUKI; KAWAGOE, 2006).

A primeira etapa para a construção da representação PAA é a fase de padronização dos

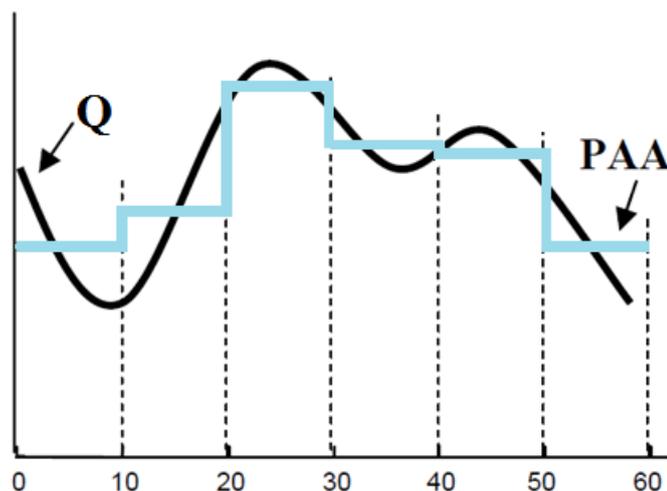


Figura 2.8: A representação PAA de uma série temporal Q. Nesse exemplo, os parâmetros PAA são $n = 60$ e $k = 6$ e cada segmento será formado pela média de 10 valores da série.

dados. Esse pré-processamento dos dados decorre do fato de que não há sentido em comparar séries temporais com amplitudes diferentes. Com isso, a padronização irá manter o formato das séries originais, mas levará os dados para a mesma amplitude relativa. Essa padronização é calculada de acordo com a Equação 2.9 (CANELAS; NEVES; HORTA, 2013). Em que, x_i é cada um dos pontos na janela da série temporal, μ_x é a média e σ_x o desvio padrão para todos os pontos presentes na janela.

$$x'_i = \frac{x_i - \mu_x}{\sigma_x} \quad (\text{Eq.2.9})$$

Dado uma série temporal de tamanho m , uma janela dessa série de tamanho n ($n \ll m$), k o número de segmentos em que irá ser reduzida a janela da série, \bar{x}_i o valor representativo do i -ésimo segmento, x_j cada um dos elementos presentes no segmento. Então, caso a relação n/k seja um valor inteiro, a operação de redução da dimensionalidade do PAA pode ser feita pela equação 2.10 (KEOGH et al., 2001):

$$\bar{x}_i = \frac{k}{n} \sum_{j=\frac{n}{k}(i-1)+1}^{\frac{n}{k}i} x_j \quad (\text{Eq.2.10})$$

No caso de a relação não resultar em um inteiro, os pontos de fronteira entre os segmentos devem contribuir para a formação do valor final de cada um dos segmentos (CANELAS; NEVES; HORTA, 2012). Nesse caso, a operação será feita como mostrado na Figura 2.9.

Como pode ser visualizado na Figura 2.9, onde existem 12 pontos e 5 segmentos, então como cada segmento deve ter o mesmo tamanho, isto significa que cada segmento deve ter a contribuição de 2,4 pontos para a média. Assim, os pontos 1 e 2 inteiramente pertencem ao segmento S1 e para completá-lo, seu valor deve contar com 40% do ponto 3. Para o segmento

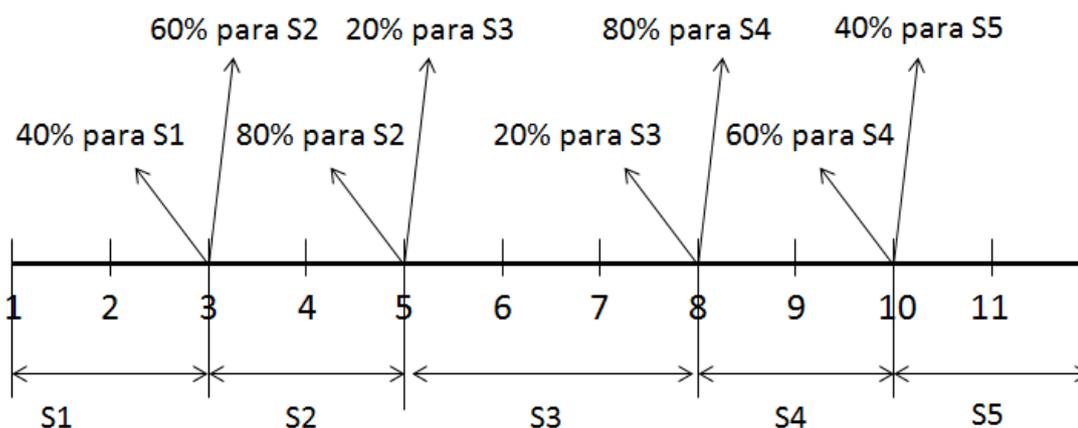


Figura 2.9: Cálculo dos valores dos segmentos PAA, quando a relação n/k não resultar em valor inteiro. Nesse exemplo, os parâmetros do PAA são $n = 12$ e $k = 5$, o que significa que cada segmento deve ter a contribuição de 2,4 pontos para a média.

S2 é formado pelos outros 60% do ponto 3, todo o ponto 4 e 80% do ponto 5, para completar os 2,4 pontos para a sua média. O resto dos segmentos são completados de acordo com esta lógica.

Entre as vantagens do PAA estão as seguintes: o algoritmo é fácil de ser implementado; é muito rápido; gera um modelo flexível; a indexação dos valores pode ser feita em tempo linear; facilidade em calcular a proximidade entre duas séries temporais, através do cálculo da distância mínima (MINDIST) entre as suas respectivas representações PAA (KEOGH et al. (2001); CHAKRABARTI et al. (2002)). Para o cálculo do MINDIST, dados \bar{Q} e \bar{C} as representações PAA de duas séries temporais, o cálculo da proximidade ocorre através da seguinte equação 2.11 (LIN et al., 2003). Este cálculo pode ser visualizado na Figura 2.10.

$$MINDIST(\bar{Q}, \bar{C}) = \sqrt{\frac{n}{k}} \sqrt{\sum_{i=1}^k (dist(\bar{q}_i, \bar{c}_i))^2} \quad (2.11)$$

Uma característica importante de MINDIST é o fato de ela representar o limite inferior da distância Euclidiana. Isso permite que o método faça a mineração dos dados sobre a representação, de forma eficiente, e produza resultados idênticos àqueles obtidos caso a operação fosse realizada sobre os dados originais (LIN et al., 2003).

Além disso, como todos os métodos de redução de dimensionalidade, PAA pode perder informações úteis para a análise de dados de séries temporais. Um exemplo é a perda de informação sobre os valores dos picos (valores extremos) presentes dentro do segmento, devido à representação pelo valor médio (LKHAGVA; SUZUKI; KAWAGOE, 2006). Por outro lado, a escolha correta dos parâmetros do PAA pode controlar essa perda de informação.

2.3.1.3 Aproximação por Valor Agregado Simbólico (SAX)

O algoritmo de Aproximação por Valor Agregado Simbólico (Symbolic Aggregate approXimation - SAX) foi proposto por LIN et al. (2003) e assim como o PAA, também é uma

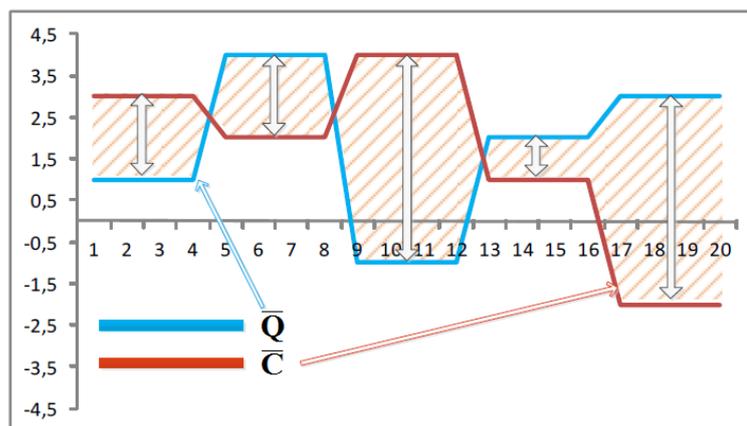


Figura 2.10: Cálculo do MINDIST entre as duas representações PAA \bar{Q} e \bar{C} .

técnica de redução de dimensionalidade dos dados. Ela pode ser dividida em duas etapas: na primeira, a série temporal é processada para a construção de sua representação PAA; então, essa representação é discretizada em uma sequência de símbolos equiprováveis, chamado Alfabeto. O número de elementos do alfabeto (α) é um parâmetro do modelo, e deve ser definido pelo usuário (LKHAGVA; SUZUKI; KAWAGOE, 2006).

Como consequência da padronização (primeira etapa no PAA), os dados a serem utilizados pelo SAX apresentam distribuição Gaussiana (LIN et al., 2003). Com isso, os intervalos equiprováveis utilizados pelos símbolos podem facilmente ser definidos através de pontos de interrupção (também denominados *breakpoints*) que produzam áreas iguais sob a curva gaussiana. Para determinar esses Pontos de Interrupção β_j , basta obtê-los em uma tabela estatística (SUN et al., 2014), uma pequena parte dela pode ser visualizada pela tabela 2.1.

Tabela 2.1: Pontos de Interrupção dividindo uma distribuição de Gauss em um número (3-5) regiões equiprováveis.

	$\alpha = 3$	$\alpha = 4$	$\alpha = 5$
β_j	-0.43	-0.67	-0.84
β_j	0.43	0	-0.25
β_j		0.67	0.25
β_j			0.84

Uma vez definidos os β_j basta utilizar a representação PAA e para cada segmento, representá-lo como o símbolo do intervalo em que ele se encontra. Dessa forma, todos os segmentos que se encontrarem abaixo do menor Pontos de Interrupção são mapeados pelo símbolo “A”; todos os maiores que o menor Pontos de Interrupção e menor que o segundo menor irão receber o símbolo “B”; e assim sucessivamente (SUN et al., 2014). A representação final da série será a concatenação de todos os símbolos, formando uma “Palavra”. A Figura 2.11 mostra um exemplo de uma representação SAX (adaptada de (LKHAGVA; SUZUKI; KAWAGOE, 2006)).

A construção da “Palavra” $\hat{c} = \hat{c}_j, \dots, \hat{c}_k$ como a representação SAX de uma série temporal

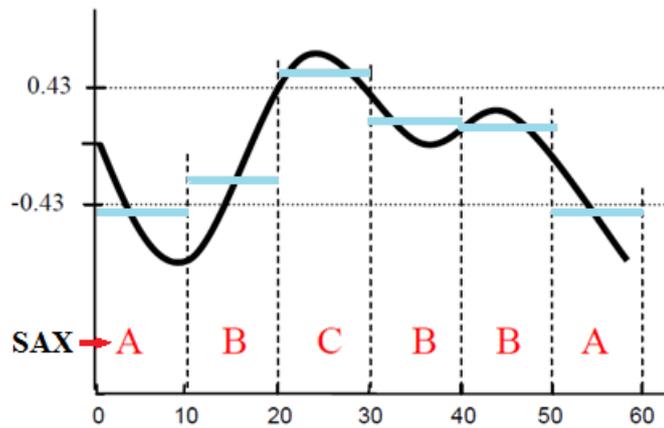


Figura 2.11: Representação SAX de uma série temporal. No exemplo acima, com $n = 60$, $k = 6$ e $\alpha = 3$, a série temporal é mapeado para a palavra ABCBBA

ou de uma janela da série temporal, de tamanho n , é definida formalmente da seguinte forma: seja α_i o i -ésimo letra do alfabeto, ou seja, $\alpha_1 = A$ e $\alpha_2 = B$. Então, a transformação de uma representação PAA \bar{c} para uma palavra \hat{c} é obtida pela equação 3.4 (LIN et al., 2003):

$$\hat{c}_i = \alpha_j, \quad sse\beta_{j-1} \leq \bar{c}_i \leq \beta_j \quad (3.4)$$

Algumas vantagens que o SAX apresenta são as seguintes (CANELAS; NEVES; HORTA, 2012): redução de dimensionalidade do problema, decorrente da primeira etapa do algoritmo (transformação da série em PAA); facilidade de visualização da representação final da série temporal, já que será formada por uma cadeia de símbolos; facilidade em calcular a proximidade entre duas séries temporais, através do cálculo da distância mínima MINDIST entre as suas respectivas palavras SAX. Dados \hat{P} e \hat{Q} como sendo palavras, o cálculo de MINDIST ocorre através da seguinte equação (LIN et al. (2003); CANELAS; NEVES; HORTA (2012)):

$$MINDIST(\hat{Q}, \hat{C}) = \sqrt{\frac{n}{k}} \sqrt{\sum_{i=1}^k (dist(\hat{q}_i, \hat{c}_i))^2} \quad (3.5)$$

Em que:

$$dist(\hat{q}_i, \hat{c}_j) = \begin{cases} 0, & |i-j| \leq 1 \\ \beta_{j-1} - \beta_i, & i < j-1 \\ \beta_{i-1} - \beta_j, & i > j-1 \end{cases}$$

2.4 Mudança de conceito (*concept drift*)

Os algoritmos projetados para aprender padrões ocultos em séries enfrentam diversos problemas, entre eles, o de manter um modelo de decisão preciso ao longo do tempo (GAMA, 2012).

Em modelos tradicionais de mineração de dados que funcionavam de forma estática, o aprendizado é realizado apenas uma vez. Nesses casos, os padrões descobertos podem se tornar obsoletos ao longo do tempo. Para solucionar esse problema, no contexto de fluxo de dados, foram propostas técnicas de processamento online de dados. Consequentemente, os modelos preditivos são retreinados a cada novo dado apresentado. Apesar da capacidade de se adaptar à evolução dos dados, estes métodos de aprendizagem podem ser computacionalmente ineficientes, por conta da atualização contínua, mesmo quando não há mudanças nas características do fluxo de dados (GAMA et al., 2014).

Nesse contexto, é importante salientar o fenômeno do “*concept drift*” ou detecção de mudança de conceito. A mudança de conceito representa uma alteração na relação entre os dados de entrada e a variável de interesse, ao longo do tempo (GAMA, 2012). Por conta disso e por ser um caso especial de fluxo de dados, séries financeiras também são afetados pelo problema *concept drift*. Modelos que incorporem mecanismos de detecção de mudança de conceito apresentam a vantagem de obterem informações sobre a dinâmica dos dados (GAMA et al., 2014). Assim, o modelo precisa ser atualizado apenas quando uma mudança de conceito for detectada, economizando recursos computacionais, quando comparado às técnicas de processamento online.

Nessa subseção serão apresentadas técnicas de detecção de mudança de conceito

2.4.1 Estado da arte: Técnicas de detecção de mudança de conceito

Diversos métodos de detecção de mudança de conceito já foram propostos. O Método de Detecção de Mudança (Drift Detection Method - DDM), proposto por GAMA et al. (2004), é um método para a detecção de mudanças na distribuição de probabilidade dos dados a partir da taxa de erro de algoritmo online. Para esta abordagem, os dados de treinamento são apresentados em sequência, quando um novo exemplo de treinamento está disponível, ele é classificado por um classificador base. A teoria estatística garante que, enquanto a distribuição for estacionária, o erro diminui. Quando a distribuição mudar, o erro aumenta.

Este modelo apresenta dois limiares: o nível de alerta e o nível de mudança. A taxa de erro p_i e o desvio padrão ($s_i = \sqrt{(p_i \cdot (1 - p_i)) / i}$) são calculados durante o treinamento e são armazenados quando $p_i + s_i$ atinge o seu menor valor, com isso, obtêm-se p_{min} e s_{min} . O nível de alerta é atingido quando $p_i + s_i \geq p_{min} + 2 \cdot s_{min}$, com isso, as novas instâncias serão armazenadas em antecipação a uma possível mudança de conceito. Quando $p_i + s_i \geq p_{min} + 3 \cdot s_{min}$, o nível de mudança é atingido, indicando uma mudança no conceito. Então, o classificador base e os valores de p_{min} e s_{min} são zerados e um novo classificador base é treinado, juntamente com novos valores para p_{min} e s_{min} , a partir dos dados armazenados do nível de alerta (GAMA et al., 2004).

O método Método Detecção Precoce de Mudança (Early Drift Detection Method - EDDM) proposto por BAENA-GARCIA et al. (2006) é semelhante ao DDM, mas, em vez de usar a taxa de erro, ele baseia-se na estimativa de distribuição das distâncias entre os erros de classificação para identificar se um desvio ocorreu. Quando não existe mudança de conceito,

o classificador base melhora suas previsões e a distância entre erros aumenta. Por outro lado, quando ocorre uma mudança de conceito, o classificador base erra mais e a distância entre erros diminui. Uma das características do EDDM é ser mais adequado para lidar com mudanças de conceito lentas e graduais.

Como no DDM, o EDDM também define dois limites: o nível de alerta e o nível de mudança de conceito. As distâncias entre dois erros de classificação p'_i e o desvio padrão ($s'_i = \sqrt{(p_i \cdot (1 - p_i)) / i}$), são calculados durante o treinamento e são armazenados quando $p'_i + 2 \cdot s'_i$ atinge o seu maior valor, com isso, obtêm-se p_{max} e s_{max} . O nível de alerta é atingido quando $((p_i + 2 \cdot s_i) / (p_{max} + 2 \cdot s_{max})) < 0.95$, mais uma vez, quando esse limite é atingido, as instâncias são armazenadas para uma possível mudança de contexto. Por sua vez, o nível de mudança é atingido quando $((p_i + 2 \cdot s_i) / (p_{max} + 2 \cdot s_{max})) < 0.90$, indicando uma provável mudança de conceito. Assim como no DDM, quando ocorre a mudança, o classificador base e os valores de p_{max} e s_{max} são zerados e um novo classificador base é treinado, juntamente com novos valores para p_{min} e s_{min} , com as instâncias armazenadas a partir do nível de alerta.

Por sua vez o método de detecção de mudança de conceito Janelamento Adaptativo (Adaptive Windowing - ADWIN) proposto por BIFET; GAVALDÀ (2007) utiliza uma abordagem de janelas deslizantes de tamanhos variáveis, que são recalculadas de forma online de acordo com a taxa de mudança observada a partir dos dados presentes nessas janelas. A detecção da mudança de conceito é feita a partir de um classificador base e seu funcionamento consiste em aumentar dinamicamente a janela (w) quando não há mudanças na distribuição dos dados, diminuindo-a quando houver. Ao ser detectada a mudança de conceito, o algoritmo busca dividir a janela w em subjanelas que apresentem médias diferentes, caso esse fato ocorra, a parte mais antiga da janela é eliminada, pois apresenta uma distribuição de dados diferente da atual.

A abordagem Classificadores Pareados (Paired Learners - PL) proposta por BACH; MALOOF (2008), como o nome sugere, usa dois classificadores: um estável e um reativo. O classificador estável classifica com base em toda a sua experiência, enquanto o reativo classifica com base em uma janela de exemplos recentes. Com isso, enquanto o método não identifica mudança no conceito, ele utiliza o classificador estável na operação de classificação, pois ele apresenta um melhor desempenho que o reativo. Por sua vez, o método usa o classificador reativo como um indicador de mudança, através da diferença de acurácia entre os dois classificadores. Com isso, caso seja identificada mudança no conceito, o algoritmo irá substituir o classificador estável pelo reativo, uma vez que, agora, este apresenta um desempenho melhor do que aquele.

A técnica Teste Estatístico de Iguais Proporções (Statistical Test of Equal Proportions - STEPD) proposta por NISHIDA; YAMAUCHI (2007) calcula a precisão do classificador base, nos casos mais recentes da janela de tempo w , comparando-a com a sua precisão geral, desde o início do processo de aprendizagem. O método assume que caso o contexto não mude, a precisão do classificador nas instâncias mais recentes presentes em w será igual a precisão global desde o início da aprendizagem; por sua vez, irá diminuir significativamente sua precisão recente caso o conceito tenha mudando. Para isso, um teste estatístico qui-quadrado é realizado e seu valor

é comparado com o percentil da distribuição normal padrão para obter o nível de significância observado. Se este valor for inferior a um nível de significância definido, a hipótese nula é rejeitada, partindo do princípio que ocorreu uma mudança de conceito.

A técnica Detecção de Mudança de Conceito Exponencial (Exponential Concept Drift Detection - ECDD), proposta por ROSS et al. (2012), usa um gráfico de médias móveis exponenciais (EMA) para monitorar a taxa de classificação incorreta de um classificador de fluxo de dados. Uma vantagem desse método é ser ele modular e, portanto, pode ser executado em paralelo com qualquer classificador subjacente para fornecer uma camada adicional de detecção de mudança de conceito. Em ECDD, esses valores EMA são calculados online, com base na precisão do classificador base e duas estimativas são comparadas: uma com mais peso em instâncias mais recentes do fluxo de dados e outro com ênfase semelhante em ambos os dados recentes e antigos. Quando a diferença entre estas duas estimativas excede um determinado limiar parametrizado, uma mudança de conceito é identificada. Os autores mencionam também um nível de aviso que é gerado quando a diferença entre estas duas estimativas está se aproximando do nível de mudança.

2.5 Otimização

2.5.1 Estado da Arte: otimização aplicada ao mercado financeiro

No problema de otimização de portfólio, os investidores buscam selecionar a melhor diversificação nos investimentos presentes no seu portfólio (carteira), segundo algum critério, como, por exemplo, menor risco. A composição dos investimentos em uma carteira depende de uma série de fatores, entre os mais importantes estão a tolerância ao risco, horizonte de investimento e valor investido (HUANG, 2012). Nesse contexto, HUANG (2012) propôs um modelo híbrido que utiliza Regressão de Vetores de Suporte (SVR) combinado com Algoritmo Genético para a seleção de ações da carteira. Nesse modelo, utiliza-se o método SVR para fornecer rankings confiáveis de ações. As ações que obtiveram melhores classificações podem, assim, ser selecionadas para formar uma carteira. Em cima deste modelo, o GA é empregado tanto para a otimização dos parâmetros do modelo, quanto para a seleção de variáveis de entrada para o modelo SVR.

Ainda quanto ao problema de otimização de portfólio, CHEN (2015) propôs um modelo utilizando Lógica Fuzzy e otimização por Colônia Artificial de Abelhas Modificado (MABC). GORGULHO; NEVES; HORTA (2011) propõem uma nova abordagem, baseada em Computação Inteligente, em especial usando Algoritmo Genético, que visa gerir uma carteira financeira por meio de indicadores de análise técnica, entre eles: Média Móvel Exponencial (EMA), Taxa de Mudança (ROC), Índice de Força Relativa (RSI), Convergência e Divergência das Médias Móveis (MACD), No Volume do Balanço (OBV). ZHU et al. (2011) apresentaram um modelo de seleção de portfólio utilizando uma variação multiobjetivo do PSO. A utilização do multiobjetivo

decorre do fato de a otimização de portfólio ser problema não-linear com mais de uma função objetivo, quais sejam o aumento do lucro e a redução do risco.

Outras abordagens no Mercado financeiro buscam a criação de *trading systems* para a automação das operações baseados em regras de decisão. TEIXEIRA; OLIVEIRA (2010) propuseram um modelo para a negociação de ações que combina uma regra de negócio, classificador baseado na regra dos k vizinhos mais próximos (KNN), indicadores de análise técnica e a otimização do algoritmo Otimização por Enxame de Partículas (PSO). Por sua vez, BRASILEIRO et al. (2013), também propuseram um trading system que utiliza indicadores técnicos e classificador baseado em KNN, mas utilizaram o algoritmo Colônia Artificial de Abelhas (ABC) como algoritmo de otimização. LI; DENG; LUO (2009) propõem um trading system que utiliza uma estratégia de negócio baseada na previsão de Picos e Vales (Turning Points Prediction), combinada com ensemble de Redes Neurais Artificiais (ANN). A previsão de Picos e Vales pode ajudar o investidor a julgar tendência do mercado e capturar oportunidades rentáveis. O Algoritmo Genético é utilizado para a otimização dos parâmetros que são utilizados pela estratégia de investimento.

O maior número de artigos na área de otimização de mercado financeiro está focado no problema de previsão de valores. Nesse contexto, ABDUAL-SALAM; ABDUL-KADER; ABDEL-WAHED (2010) apresentam um estudo comparativo sobre o uso dos algoritmos de otimização Evolução Diferencial (DE) e PSO aplicados à rede neural feed-forward usadas na previsão das cotações diárias do mercado de ações. PULIDO; MELIN; CASTILLO (2014) apresentam um método híbrido baseado em Particle Swarm Optimization (PSO) para a construção de ensemble de redes neurais artificiais (ANN). Esses ensembles são combinados com lógica fuzzy para a previsão de séries temporais da Bolsa de Valores Mexicana. GUO-QIANG (2011) propôs um modelo para a previsão do preço das ações baseado em Regressão de Vetor de Suporte (SVR), para prever o preço de fechamento das ações no terceiro dia futuro através do estudo sobre o preço das ações de alguns dias anteriores. A ideia de prever o preço no terceiro dia é dar ao investidor um maior tempo para tomar suas decisões. Essa abordagem utiliza PSO para otimizar os parâmetros do SVR. EVANS; PAPPAS; XHAFA (2013) apresentam um modelo de previsão e tomada de decisão baseada em Redes Neurais Artificiais e Algoritmo Genético. Essa abordagem é utilizada para previsão de preço intraday do Mercado Cambial (FOREX) para seguintes conversões de moedas: Libra Esterlina para Dólar Americano; Euro para Libra Esterlina e Euro para Dólar Americano.

XIONG et al. (2015) propuseram uma nova abordagem de previsão de séries temporais baseada em redes neurais FCRBFNN (fully complex-valued radial basis function neural networks). Os algoritmos PSO e Discrete PSO (DPSO) são utilizados de forma conjunta para otimizar a estrutura e o conjunto de parâmetros do modelo. KAZEM et al. (2013) propõem um modelo de previsão do preço de mercado de ações com base nas técnicas de Regressão de vetores de suporte (SVR) combinada com a otimização do Algoritmo dos Vagalumes baseado em Caos (Chaos-based Firefly Algorithm - CFA). Para esse modelo, inicialmente o CFA é utilizado para

otimizar os hiperparâmetros do SVR que posteriormente é utilizado para a previsão das cotações. ASADI et al. (2012) apresentam um modelo híbrido de previsão em bolsa de valores. Essa abordagem utiliza uma combinação de métodos de pré-processamento de dados, Algoritmos Genéticos (GA) e o algoritmo de Levenberg-Marquardt (LM) para aprendizagem redes neurais feed forward. As técnicas de pré-processamento utilizadas são transformação de dados e seleção de variáveis de entrada. Por sua vez o Algoritmo Genético e LM são utilizados para refinar os pesos iniciais da rede neural.

Por fim, algoritmos de otimização ainda são combinados com técnicas de mineração de dados para serem aplicados ao Mercado financeiro. MAJHI et al. (2008) propõem uma abordagem de mineração de dados que utiliza uma variação da Otimização por Enxame de Partículas (Clonal Particle Swarm Optimization - CPSO) e técnicas de PSO para desenvolver modelos adaptativos para previsão de cotações dos preços de ações no curto e no longo prazo da S&P 500 e índices de ações Dow Jones, nos Estados Unidos. A estrutura básica dos modelos é um combinador linear adaptativo, baseado em indicadores de análise técnica, cujos pesos são atualizados de forma iterativa por regras de aprendizagem que utilizam o PSO e o CPSO. MAJHI et al. (2009) apresenta uma abordagem de mineração que usa de técnicas de Otimização de Forrageio Bacteriana (bacterial foraging optimization (BFO) e sua versão adaptativa, o ABFO, para construção de modelos utilizados na previsão de vários índices de ações. A estrutura utilizada para combinação desses modelos de previsão é um combinador linear simples. Os pesos de conexão dos modelos baseados combinador linear adaptativo são otimizados com ABFO e BFO, minimizando o erro quadrado médio (MSE).

CANELAS; NEVES; HORTA (2012) propuseram uma nova abordagem combinando a técnica simbólica SAX, juntamente com um kernel de otimização baseada no Algoritmos Genéticos. A representação SAX é usada para descrever a série temporal financeira, de modo que, padrões relevantes possam ser identificados de forma eficiente. O Algoritmo Genético é aqui utilizado para identificar os padrões mais relevantes e gerar regras de investimento. Posteriormente, em CANELAS; NEVES; HORTA (2013), o modelo proposto foi aprimorado. Neste, foram propostas novas regras de negócio, nova estrutura de indivíduo utilizado pelo GA além de as operações ocorrerem tanto na posição comprando quanto na posição vendido.

2.5.2 Algoritmo Genético

O Algoritmo Genético (GA) é um método de otimização que simula o processo de evolução natural para buscar a melhor solução. As principais operações que são feitas ao longo das gerações do algoritmo são a seleção, o cruzamento (crossover) e a mutação, podendo ser utilizada ainda a estratégia do elitismo (HASSAN et al., 2005). No GA, cada solução candidata do problema é chamada de indivíduo (cuja estrutura é chamada de cromossomo) que são, inicialmente, gerados de forma aleatória para a criação da população inicial. Em seguida, esses indivíduos são colocados no ambiente de problema, e aqueles que melhor se adaptarem

ao ambiente, de acordo com o princípio da sobrevivência do mais apto, pode ser selecionado e copiado (elitismo), serem selecionados para fazer o cruzamento com outros cromossomos (crossover), ou sofrerem mutação de seus elementos, para, com isso, formar a próxima geração da população. A cada geração, os indivíduos são avaliados de acordo com o seu grau de aptidão (função fitness), de modo que o indivíduo com alta aptidão podem participar da operação de herança e o indivíduo com baixa aptidão será eliminado. O algoritmo é executado até que seja atingido o critério de parada que, geralmente, ocorre quando algum indivíduo da população atingiu o grau de aptidão (fitness) desejado, caso este seja conhecido a priori, ou porque foi atingido o número máximo de gerações permitido. Ao final, o indivíduo, que melhor se adaptar ao ambiente será a solução ideal para o problema (LI et al., 2010).

A operação de seleção é aquela que seleciona os indivíduos de maiores aptidão para geração da população filha através do cruzamento e da mutação, objetivando dirigir o GA para as melhores regiões do espaço de busca. A seleção no GA pode ser implementada de diversas maneiras, entre elas estão as seguintes: seleção de roleta, em que indivíduos com maior aptidão têm mais chance de ser selecionado; e seleção por competição, em que se escolhem n indivíduos aleatoriamente da população, em regra utilizam-se dois, e o melhor é selecionado. Independentemente do método de seleção utilizado, uma estratégia elitista é muitas vezes utilizada, em que o indivíduo com maior aptidão sempre será copiado para a próxima geração (EBERHART; SHI, 1998).

Por sua vez, a operação de cruzamento ocorre através da combinação dos pais selecionados aleatoriamente da população para a produção dos filhos. Os efeitos dessa operação, em regra, variam ao longo das gerações. Isto é, no início, como os indivíduos da população são gerados aleatoriamente, a população filha tem uma capacidade maior de exploração do espaço de busca problema. Por sua vez, em iterações mais avançadas, como os indivíduos já irão convergir para um ponto, apresentarão estruturas semelhantes, e essa operação de cruzamento já não vai apresentar efeito significativo quanto a capacidade de exploração. Por conta disso, versões mais atuais do GA estão implementando uma probabilidade de cruzamento que varia ao longo das gerações, muitas vezes começando a com probabilidade relativamente grande, e terminando com um menor (EBERHART; SHI, 1998).

Algoritmo 1: Pseudocódigo GA

```
1 início
2   Gerar a população inicial
3   Avaliar o fitness de cada indivíduo da população.
4   repita
5     Selecionar os indivíduos mais aptos
6     Criar novos indivíduos aplicando os operadores cruzamento, elitismo e
       mutação
7     Armazenar os novos indivíduos em uma nova população
8     Avaliar cada cromossomo da nova população.
9   até critério de parada seja atingido;
10 fim
```

Por fim, a operação de mutação ocorre através da mudança de valores individuais do cromossomo do indivíduo. Os efeitos da mutação tendem a ser o oposto da operação de cruzamento. Com isso, a mutação tende a ter menor impacto nas gerações iniciais e maior nas finais. Isso ocorre porque, como no início os indivíduos são gerados aleatoriamente, a troca de um valor individual não irá alterar a aptidão do indivíduo de forma drástica do que aquelas causadas no indivíduo quando a população tenha convergido. Se a taxa de mutação variar ao longo das iterações, é comum utilizar valores menores no início, aumentando-a próximo ao fim (EBERHART; SHI, 1998). O pseudocódigo do GA é apresentado no Algoritmo 1.

2.5.3 Otimização Por Enxame de Partículas

A Otimização por Enxame de Partículas (Particle Swarm Optimization - PSO) foi proposto por KENNEDY; EBERHART (1995). Este algoritmo é um método de otimização estocástica usado para encontrar o ótimo global de uma função objetivo através da simulação do comportamento de um bando de aves e de um cardume de peixes. Comparado com outros métodos evolutivos, PSO tem as seguintes vantagens: o algoritmo apresenta uma estrutura simples; ajuste de parâmetros simples; velocidade de convergência rápida; possui uma boa *trade-off* entre exploração e de exploração. Esta deve ser entendida como a capacidade de refino das soluções encontradas, enquanto aquela é a capacidade de explorar ao máximo o espaço de busca (EBERHART; SHI, 1998).

No PSO, inicialmente, é gerado aleatoriamente um “enxame de partículas” (que representa o conjunto de possíveis soluções) num espaço de busca multidimensional. Dentro desse espaço, cada indivíduo da população é denominado partícula e apresenta dois valores associados, sua posição e sua velocidade, que são modificados a cada iteração. A posição representa o seu lugar no espaço. Por sua vez, a velocidade é um valor resultante da influência dos seguintes fatores que se somam: inércia, componente cognitivo e componente social. A inércia ajuda a partícula a manter a direção que vinha seguindo. O componente cognitivo atua na atração da partícula na direção da sua melhor posição até o momento. Já o componente social atrai a partícula na direção da melhor solução encontrado por toda a população, até a atual iteração. Dessa forma, cada partícula ajusta sua posição de acordo com sua experiência e a dos demais indivíduos da população, combinando, assim, uma de busca local com uma busca global. Por fim, a qualidade das posições obtidas pelas partículas é calculada a partir de uma função de aptidão. (HASSAN et al., 2005). O pseudocódigo do PSO é apresentado no Algoritmo 1.

Para a definição formal do algoritmo, seja um espaço de busca com D dimensões e um enxame com N partículas, a i -ésima partícula do enxame pode ser representada por um vetor D -dimensional $X_i = [x_{i1}, x_{i2}, \dots, x_{iD}]$, sendo esse o vetor correspondente a posição da partícula no espaço; a velocidade da partícula consiste em $V_i = [v_{i1}, v_{i2}, \dots, v_{iD}]$; a melhor posição que a partícula como $P_i = [p_{i1}, p_{i2}, \dots, p_{iD}]$ e a melhor posição obtida pelo enxame como $P_g = [p_{g1}, p_{g2}, \dots, p_{gD}]$. Então, para cada geração, a atualização dos valores da velocidade e da

Algoritmo 2: Pseudocódigo PSO

```

1 início
2   Gerar uma população inicial de partículas:  $P$ 
3   Inicializar aleatoriamente a posição inicial ( $x$ ) e a velocidade inicial ( $v$ ) de cada
   partícula  $i$  de  $P$ .
4   para cada partícula  $i$  de  $P$  faça
5     |   Calcular sua aptidão  $f_i$  através da função fitness escolhida
6     |   Calcular a melhor posição da partícula  $i$  até o momento:  $p_{ibest}$ 
7   fim
8   Selecionar a partícula com a melhor aptidão de toda a população:  $p_{gbest}$ 
9   para cada partícula  $i$  de  $P$  faça
10    |   Atualizar a velocidade da partícula: Eq. 2.12
11    |   Atualizar a posição da partícula: Eq. 2.13
12  fim
13  se critério de parada não for alcançada então
14    |   Retorne à linha 4.
15 fim

```

posição para cada partícula do enxame, é, respectivamente, definida pelas equações 2.12 e 2.13 (ZHANG; XIONG; ZHANG, 2013):

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot rand \cdot (p_i - x_i(t)) + c_2 \cdot Rand \cdot (p_g - x_i(t)) \quad (Eq.2.12)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (Eq.2.13)$$

Em que, c_1 e c_2 representam fatores de aceleração e são constantes positivas, $rand$ e $Rand$ são duas variáveis randômicas com distribuição uniforme dentro de intervalo $[0,1]$, e w é a ponderação de inércia. Na equação da velocidade o primeiro fator representa a inércia, o segundo fator o componente cognitivo e o terceiro fator o componente social (ZHANG; XIONG; ZHANG, 2013).

2.5.4 Otimização por Enxame de Partículas Auto Adaptativa

Ao longo da utilização do PSO, as pesquisas demonstraram que os três parâmetros w , c_1 e c_2 têm um impacto significativo no desempenho do algoritmo. Um peso maior de inércia facilita a exploração global (exploração em busca de novas áreas), enquanto um peso de inércia menor tende a facilitar a exploração local para o refino da solução na área de pesquisa atual (exploração). Quanto aos valores dos componentes cognitivo e social, maiores valores para c_1 dão um maior peso para a busca local, enquanto maiores valores para c_2 dão maior relevância para a busca global (HASSAN et al., 2005).

Por isso, ZHANG; XIONG; ZHANG (2013) propuseram o algoritmo Otimização por Enxame de Partículas Auto Adaptativa (Improved self-adaptive particle swarm optimization - IDPSO). Nessa variação do PSO, o próprio algoritmo ajusta w , c_1 e c_2 dinamicamente através do

valor do calculo da função de detecção $\varphi(t)$ definida, levando em consideração tanto a busca global quanto a busca local. A função de detecção é definida pela equação 2.14:

$$\varphi(t) = |(p_g - x_i(t-1))|/|(p_i - x_i(t-1))| \quad (Eq.2.14)$$

Em que, $|(p_g - x_i(t-1))|$ representa a distância Euclidiana entre a melhor posição encontrada pelo enxame de partículas e a posição anterior da partícula i ; e $|(p_i - x_i(t-1))|$ representa a distância Euclidiana entre a melhor posição encontrada pela partícula i e sua a posição anterior.

Os valores dos dois fatores de aceleração c_1 e c_2 são modificados dinamicamente de acordo com a função de detecção $\varphi(t)$ e a variação do peso de inércia é baseada tanto nos valores de $\varphi(t)$ quanto numa função sigmoide variável para encontrar a melhor solução. A atualização das variáveis utilizadas pelo algoritmo pode ser expressa pelas equações 2.15, 2.16, 2.17, 2.18 e 2.19:

$$w(t) = \frac{w_{inicial} - w_{final}}{1 + e^{\varphi(t) \cdot (t - ((1 + \ln(\varphi(t))) \cdot K_{max}) / \mu)}} + w_{final} \quad (Eq.2.15)$$

$$c_1 = c_1 \cdot \varphi(t)^{-1} \quad (Eq.2.16)$$

$$c_2 = c_2 \cdot \varphi(t) \quad (Eq.2.17)$$

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot rand \cdot (p_i - x_i(t)) + c_2 \cdot Rand \cdot (p_g - x_i(t)) \quad (Eq.2.18)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (Eq.2.19)$$

Em que, $w_{inicial}$ e w_{final} representam, respectivamente, os valor inicial de final da inércia w (valores no intervalo $0 < w < 2$). K_{max} representa o número máximo de iterações utilizadas no algoritmo, t representa a iteração atual do algoritmo, $\varphi(t)$ representa a função de detecção e μ representa um fator de ajuste. Os novos componentes apresentados pelo IDPSO devem ser entendidos da seguinte forma: o uso da função de detecção $\varphi(t)$ dá ao algoritmo a capacidade de realizar a transição suave de “exploração” para “exploração” de forma adaptativa com base na função sigmoial.

- Nas iterações iniciais: $\varphi(t) \geq 1$, ou seja, $|(p_g - x_i(t-1))| > |(p_i - x_i(t-1))|$:
 - O algoritmo enfatiza a busca global e fortalece a capacidade de “exploração”.
 - O peso do componente social (eq. 2.17) aumenta para melhorar a partilha de informação e a cooperação das partículas, de modo que o valor de c_1 é diminuído e o valor de c_2 é aumentado.
- Nas iterações mais avançadas: $\varphi(t) < 1$, ou seja, $|(p_g - x_i(t-1))| < |(p_i - x_i(t-1))|$:
 - O algoritmo dá ênfase à capacidade de busca local e fortalece a capacidade de “exploração”.

- O peso do componente cognitivo (eq. 2.16) aumenta para reforçar a própria influência da partícula, de modo que o valor de c_1 é aumentado e o valor de c_2 é diminuído.

Com isso, o IDPSO apresenta as seguintes vantagens: apresenta todas as características do PSO e ainda aprimora sua capacidade de exploração e de exploração, além de o IDPSO não ser dependente dos parâmetros w , c_1 e c_2 (ZHANG; XIONG; ZHANG, 2013). O pseudocódigo do IDPSO está representado no Algoritmo 2.

Algoritmo 3: Pseudocódigo IDPSO

```

1 início
2   Gerar uma população inicial de partículas:  $P$ 
3   Inicializar aleatoriamente a posição inicial ( $x$ ) e a velocidade inicial ( $v$ ) de cada
   partícula  $i$  de  $P$ .
4   para cada partícula  $i$  de  $P$  faça
5     Calcular sua aptidão  $f_i$  através da função fitness escolhida
6     Calcular a melhor posição da partícula  $i$  até o momento:  $p_{ibest}$ 
7   fim
8   Selecionar a partícula com a melhor aptidão de toda a população:  $p_{gbest}$ 
9   para cada partícula  $i$  de  $P$  faça
10    Calcular a função detecção  $\varphi(t)$ . Eq. 2.14
11    Atualizar o valor da inércia: Eq. 2.15
12    Atualizar das variáveis  $c_1$  e  $c_2$ : Eq. 2.16 e Eq. 2.17
13    Atualizar a velocidade da partícula: Eq. 2.18
14    Atualizar a posição da partícula: Eq. 2.19
15  fim
16  se critério de parada não for alcançada então
17    Retorne à linha 4.
18 fim

```

2.6 Considerações finais

Nesse capítulo, foram apresentados conceitos básicos para o entendimento sobre o uso de padrões gráficos em mercados financeiros, séries temporais, mineração de séries temporais, detecção de mudança de conceito e otimização.

Em relação aos padrões gráficos, foram apresentados alguns padrões que já são utilizados na prática no mercado financeiro. Ainda foi dado destaque às dificuldades enfrentadas na descoberta desses padrões sem o uso de sistemas computacionais, cenário no qual o presente estudo se insere.

Na sequência, com o objetivo de contextualizar o presente trabalho foi apresentado um breve estado da arte de uso de técnicas de representação de séries temporais aplicadas ao mercado financeiro e, ainda, um maior detalhamento sobre as técnicas PAA e SAX, que são, respectivamente, a técnica utilizada no presente trabalho e a abordagem de comparação. No

entanto, a maior parte dessas técnicas de mineração de dados tradicionalmente utilizadas são estáticas, não levando em consideração mudanças que ocorram ao longo do uso do sistema. Isso pode gerar dificuldades no uso prático dessas abordagens já que no mundo real podem ocorrer tais mudanças. Diante disso, o presente desenvolveu um método dinâmico baseado na detecção de mudança de conceito, que possa se adequar às mudanças nas características do ambiente ao longo de seu uso.

Na detecção de mudança de conceito, foram apresentadas técnicas que já são utilizadas e difundidas na literatura. Como as metodologias apresentadas necessitam de classificadores para a detecção de mudança de conceito e o nosso modelo não utiliza classificadores, há a necessidade de desenvolvimento de uma métrica que possa se adequar ao presente trabalho.

Por fim, na seção de otimização foi listado uma breve revisão da literatura da aplicação dos algoritmos dessa área em diversos contextos presentes no mercado financeiro. O conhecimento acerca das características dos ambientes de busca e dos algoritmos de otimização são de fundamental importância para os métodos possam obter ser capazes de tirar o melhor proveito das informações. Além disso, na seção ainda foram detalhadas os algoritmos GA, PSO e IDPSO, este último como o algoritmo utilizado pela versão final do modelo proposto e os demais servindo para fins de comparação.

3

Método Proposto

3.1 Introdução

Neste capítulo é apresentado um novo método para descoberta de padrões em séries temporais financeiras, o PAA-IDPSO-CD. Essa abordagem é formada a partir da combinação do método de representação de séries temporais Aproximação por Valor Agregado de Segmento (PAA) com o algoritmo de otimização Otimização por Enxame de Partículas Auto Adaptativa (IDPSO), junto a uma métrica de detecção de mudança de conceito proposta, responsável por identificar os momentos de retreino, em que os padrões e os parâmetros do modelo serão recalculados, durante o uso do método (ou fase de teste). Os padrões encontrados são, então, utilizados junto a regras de decisão propostas com o objetivo de servir como um sistema de suporte de decisão responsável por automatizar as operações feitas no mercado de ações. É proposta também uma técnica para a redução da variância do modelo.

O método proposto utiliza tanto a ideia de descoberta quanto a de reconhecimento de padrões. A descoberta é realizada na fase de treinamento, em que o modelo busca padrões desconhecidos, dentro dos dados da série temporal e também otimiza parâmetros da regra de negócios, como o número máximo de dias que o sistema deve manter a ação comprada. Por sua vez, na fase de teste, o modelo realiza a operação de reconhecimento, pois busca padrões iguais ou semelhantes ao já conhecido (FU, 2011).

A figura 3.1 mostra a evolução do modelo até chegar a sua forma final. A versão inicial não apresenta retreino na fase de teste e opera apenas em posições comprado. Por sua vez, a segunda versão também não apresenta retreino, mas as regras de decisão opera nas posições comprado e vendido. Por fim, o modelo final proposto possui o mecanismo de retreino durante a fase de teste e pode operar tanto em posições comprado e vendido quanto apenas em comprado.

3.2 PAA-IDPSO operando apenas em posição comprado

O algoritmo PAA-IDPSO (Aproximação por Valor Agregado de Segmento - Otimização por Enxame de Partículas Auto Adaptativa) combina PAA, técnica de representação de dados

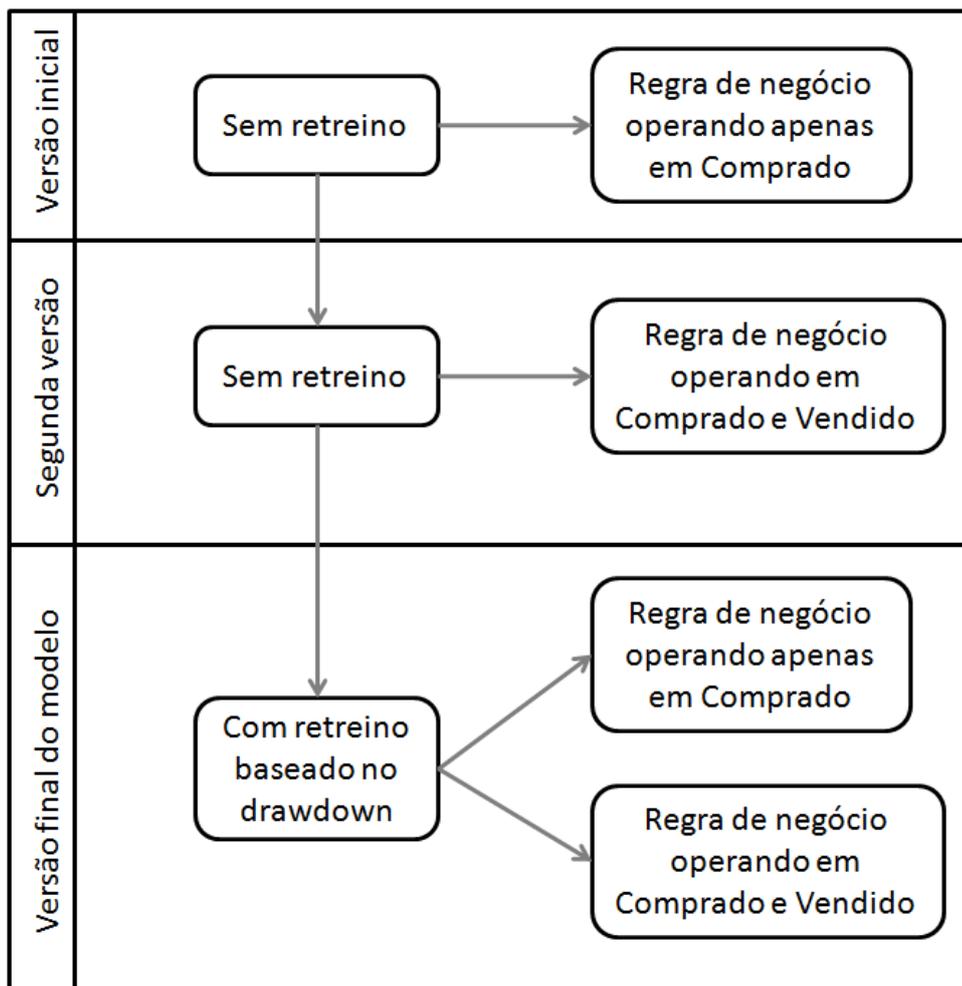


Figura 3.1: Evolução do modelo proposto até chegar a sua forma final.

utilizada no contexto de mineração de dados de séries temporais financeiras, com o algoritmo IDPSO. Essa combinação é utilizada na descoberta dos melhores padrões que representam uma série temporal financeira que serão utilizados por uma regra de negócio proposta, com o objetivo de servir como sistema de suporte de decisão capaz de automatizar as operações feitas no mercado de ações. O algoritmo é dividido em duas etapas: fase de treinamento e uso do sistema (ou fase de teste). No treinamento são definidos os melhores parâmetros a serem utilizados pela regra de negócio e o padrão que melhor representa a série temporal. Por sua vez, na etapa de teste o padrão já definido será utilizado para a criação de ordens de compra e venda pela regra de negócio.

O presente estudo optou pelo uso da abordagem Aproximação por Valor Agregado de Segmento (PAA), porque ela é uma técnica de representação dos dados utilizada para a redução da dimensionalidade de dados e, conseqüentemente, diminuição do espaço de armazenamento e do tempo de processamento computacional. Além disso, as representações PAA podem ser utilizadas na descoberta de padrões representativos nas séries temporais (FU, 2011). Outras vantagens dessa técnica são as seguintes: 1) ser rápida, simples e surpreendentemente competitiva quando comparada a outras técnicas de transformações mais sofisticadas, 2) poder ser aplicado

diretamente sem processamento adicional dos dados originais disponíveis e 3) o fato de cada segmento PAA ter o mesmo comprimento facilita indexação desta representação (CHAKRABARTI et al., 2002). Por fim, o PAA, quando comparado à técnica SAX, apresenta como vantagens o fato de ser mais rápida e necessitar de menos parâmetros. Uma vez que, enquanto os parâmetros do PAA são apenas o tamanho da janela (n) e a quantidade de segmentos (k), o SAX precisa de um parâmetro adicional, que é o tamanho do alfabeto (α). Esta é uma vantagem muito importante do PAA em relação ao SAX, pois a escolha correta de α é fundamental para o desempenho SAX. Além do fato de, diferentemente do SAX, o PAA utilizar valores contínuos em sua representação (LIN et al., 2003).

Como ponto negativo do PAA, assim como nas demais técnicas de redução de dimensionalidade, é a possibilidade de haver perda de informações úteis para a análise das séries temporais, como, por exemplo, perda de informações a respeito de valores extremos, da variância e/ou da dispersão dos dados da série. Mesmo assim, as vantagens que a abordagem traz em relação à redução de tempo de processamento e de espaço de armazenamento podem compensar essa perda. Além disso, espera-se que a combinação do PAA com a otimização do IDPSO e da detecção de mudança de conceito possam, ainda mais, reduzir os efeitos da perda de informação a respeito das séries temporais.

O fato de o PAA trabalhar com variáveis contínuas foi decisivo para a opção pelo uso do PSO. Uma vez que, de acordo com a literatura este algoritmo apresenta boa eficiência computacional, quando utilizado para resolver problemas com variáveis contínuas (HASSAN et al., 2005). Além disso, o PSO ainda tem o diferencial de ser um algoritmo com estrutura simples, ajuste de parâmetros simples, velocidade de convergência rápida e capacidade de combinação de busca local com busca global. ((LI et al., 2010); (EBERHART; SHI, 1998)). Por sua vez, a opção pelo IDPSO se deu, principalmente, por causa da sua melhoria nas capacidades de exploração e de exploração do PSO padrão, além de não ser dependente do parâmetros inércia w , aceleração do componente cognitivo c_1 e aceleração do componente social c_2 ((ZHANG; XIONG; ZHANG, 2013)).

Para o modelo proposto, os parâmetros do PAA serão definidos pelo usuário. Além disso, a técnica utilizada para percorrer a série temporal financeira será a técnica de janela deslizante com *lag* um, ou seja, a cada iteração do algoritmo adianta-se apenas um dia da série temporal (cotações diárias das ações).

Para o IDPSO, a estrutura da partícula será semelhante à estrutura do indivíduo, apresentado pela SAX-GA (CANELAS; NEVES; HORTA, 2012). A Figura 3.2 mostra a representação de uma partícula.

Como pode ser observado na Figura 3.2, a definição do número D de dimensões das partículas depende do número de segmentos (k) da representação PAA:

- Uma dimensão representa “Distância para Compra”;
- Uma dimensão representa “Distância para Venda”;

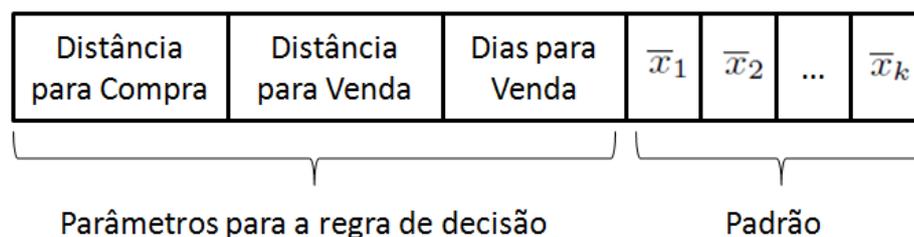


Figura 3.2: A representação da estrutura da partícula com regra de negócio operando apenas em posições comprado.

- Uma dimensão representa “Dias para Venda”;
- k dimensões para a representação PAA do melhor padrão representativo da série temporal a ser utilizado em ambas as decisões de entrada e saída de posição comprado:
 - Representação do padrão único que será utilizado para as decisões de entrada e saída de posição comprado: $PAA = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k]$

A função de aptidão que o IDPSO irá otimizar é o total dos rendimentos produzidos pela estratégia de investimento definido pelo padrão e aplicação das regras de decisão associadas a ele. Como a representação PAA é usada, os padrões são sequências de valores PAA (média dos valores no segmento) e a função MINDIST para PAA é usada para calcular a similaridade entre o padrão descoberto e a representação PAA para uma janela das séries temporais financeiras. Esta distância define o quão próxima a representação PAA da janela atual da série deve estar do padrão descoberto para acionar uma ordem de COMPRA, ou quão longe elas estão para disparar uma ordem de VENDA. Além disso, as operações são feita ação à ação, isto é, todo o saldo é aplicado ao modelo que opera apenas sobre uma ação, não sendo utilizada a ideia de portfólio de ações.

Para esta primeira versão, a regra de negócio do PAA-IDPSO opera apenas na posição de comprado (*long position*), isto é, compra de uma ação com a expectativa de que o seu preço irá subir em valor, com a expectativa de lucro com a venda da ação (TUNG; QUEK, 2011). Então, no final da etapa de treinamento, o algoritmo retorna os melhores valores para os parâmetros obtidos pela melhor partícula do enxame. “Distância para Compra” representa o quão próximas duas representações devem estar para provocar uma compra de ações. “Distância para Venda” o quão distantes duas representações devem estar para se justificar uma venda, caso haja alguma ação comprada. “Dias para Venda” representa o número máximo de dias em que o sistema irá operar com as ações compradas. Após esse número de dias, o sistema vende as ações independentemente da comparação com o padrão descoberto no treinamento. Por fim, o algoritmo retorna também a melhor representação PAA encontrada pelo IDPSO, como melhor padrão, para usá-lo na etapa de teste.

Na fase de teste, o algoritmo calcula a similaridade entre o melhor padrão e as representações PAA das janelas deslizantes da série original (subconjunto de teste). Então, essa distância é

Algoritmo 4: Regra de negócio para operar apenas em posições comprado

Dados:
 Dist_Compra = Distância para compra definida pela melhor partícula
 Dist_Venda = Distância para venda definida pela melhor partícula
 DIAS = Limite da quantidade máxima de dias para operar comprado definido pela melhor partícula
 PADRÃO = representação PAA definido pela melhor partícula
 Janela = janela da série temporal original retornada pela janela deslizante

```

1 início
2   se ( $MINDIST(PADRÃO, PAA(Janela)) < Dist\_Compra$ ) então
3     | Gera ordem de COMPRA
4   se COMPRADO então
5     | se ( $MINDIST(PADRÃO, PAA(Janela)) > Dist\_Venda$ ) então
6     |   | Gera ordem de VENDA
7     | se Tempo de ação comprada > DIAS: então
8     |   | Gera ordem de VENDA.
9 fim
```

comparada com os parâmetros retornados pela fase de treinamento para decidir se deve comprar as ações ou se deve vendê-las. A regra de decisão também calcula o número de dias em que ação ficou comprada para vendê-la, independentemente da distância calculada, quando atinge o limite estabelecido nos “Dias para Venda”, calculado na fase de treinamento.

As regras de decisão utilizadas pelo PAA-IDPSO na regra de negócio operando apenas em posição comprado são definidas no Algoritmo 3. Para uma melhor visualização, as regras também são apresentadas na figura 3.3.

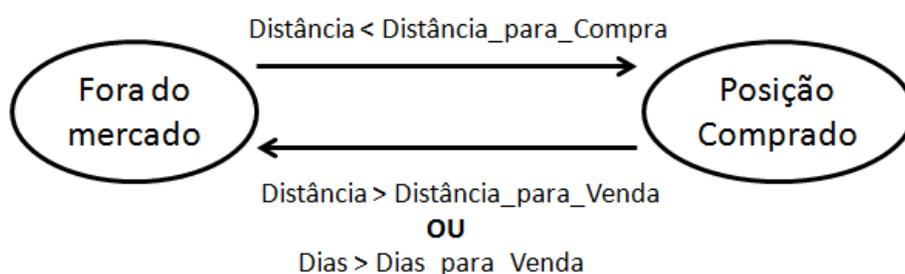


Figura 3.3: Diagrama representando a regra de negócio operando apenas em posições comprado.

3.3 PAA-IDPSO operando nas posições comprado e vendido

Nessa segunda versão, a regra de negócio do PAA-IDPSO opera tanto na posição de comprado (*long position*), quanto na posição de vendido (*short position*). Esta última opera em

sentido contrário àquela, isto é, vende-se uma ação emprestada com a expectativa de que a ação cairá em valor (TUNG; QUEK, 2011).

Em vez de apenas encontrar um padrão para a escolha do momento para investir, esta nova estratégia vai encontrar um padrão para entrar na posição de comprado, um padrão para sair desta, um padrão para entrar na posição vendido e um padrão para sair da posição vendido. Além disso, como a nova estrutura opera tanto em posições comprado quanto em vendido, então há a necessidade de um limite máximo de dias para cada uma dessas estratégias.

Então, basicamente a abordagem descrita nessa seção representa uma ampliação da primeira versão. Dessa forma, a estrutura da partícula apresentará uma forma diferente daquela apresentada pela primeira versão. A Figura 3.4 mostra a nova representação de uma partícula.

Distância para entrar Comprado	Dias para Venda			x_1^{EL}	x_2^{EL}	...	x_k^{EL}
Distância para sair Comprado	x_1^{SL}	x_2^{SL}	...	x_k^{SL}			
Distância para entrar Vendido	Dias para Compra			x_1^{ES}	x_2^{ES}	...	x_k^{ES}
Distância para sair Vendido	x_1^{SS}	x_2^{SS}	...	x_k^{SS}			

Figura 3.4: A representação da estrutura da partícula com regra de negócio operando em posições comprado e vendido.

Como pode ser observado na Figura 3.4, a definição do número D de dimensões de partículas também depende do número de segmentos (k) da representação PAA:

- Uma dimensão representa “Distância para entrar na posição Comprado”;
- Uma dimensão representa “Dias para Venda”;
- k dimensões para a representação PAA do melhor padrão de entrada em posição comprado a ser encontrado
 - Representação do padrão de entrada em posição comprado: $PAA = [x_1^{EL}, x_2^{EL}, \dots, x_k^{EL}]$
- Uma dimensão representa “Distância para sair da posição Comprado”;
- k dimensões para a representação PAA do padrão de saída em posição comprado a ser encontrado
 - Representação do padrão de saída da posição comprado: $PAA = [x_1^{SL}, x_2^{SL}, \dots, x_k^{SL}]$
- Uma dimensão representa “Distância para entrar na posição Vendido”;

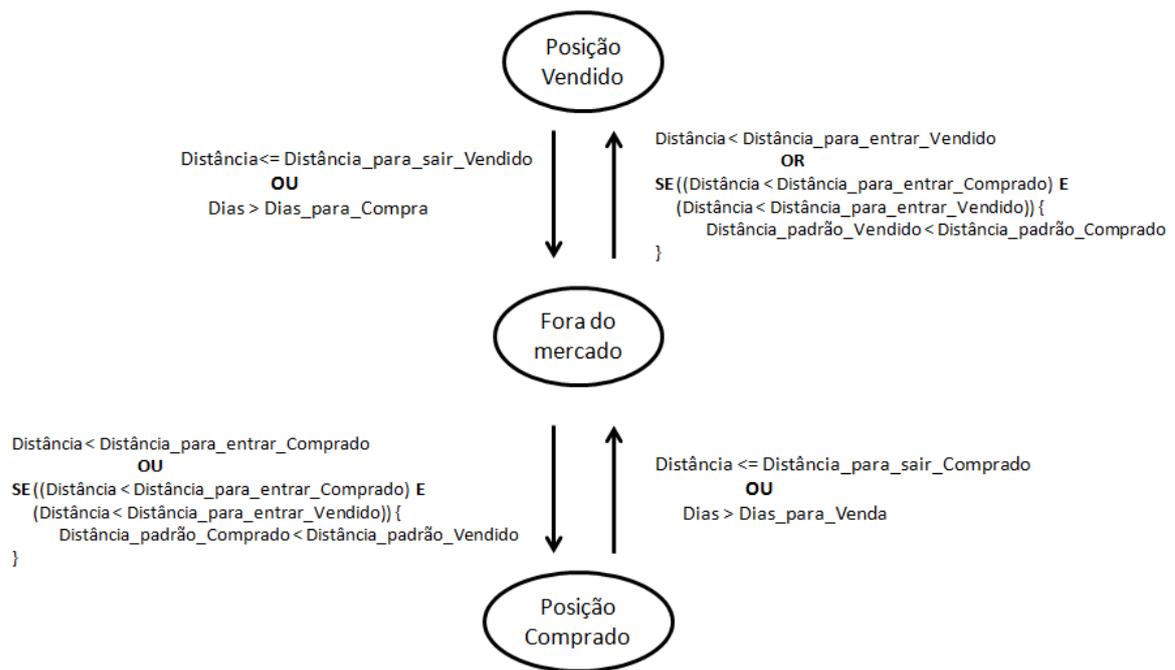


Figura 3.5: Diagrama representando a regra de negócio para operar em posições comprado e vendido.

- Uma dimensão representa “Dias para Compra”;
- k dimensões para a representação PAA do padrão de entrada em posição vendido a ser encontrado
 - Representação do padrão de entrada na posição vendido: $PAA = [x_1^{ES}, x_2^{ES}, \dots, x_k^{ES}]$
- Uma dimensão representa “Distância para sair da posição Vendido”;
- k dimensões para a representação PAA do padrão de saída em posição vendido a ser encontrado
 - Representação padrão de saída da posição vendido: $PAA = [x_1^{SS}, x_2^{SS}, \dots, x_k^{SS}]$

As regras de decisão utilizadas pelo PAA-IDPSO operando nas posições comprado e vendido são definidas no Algoritmo 4. Para uma melhor visualização, as regras também são apresentadas na figura 3.5.

Nessa nova abordagem, ao final da etapa de treinamento, o algoritmo também retornará o vetor contendo os melhores valores para os parâmetros obtidos pela melhor partícula do enxame, isto é, o conjunto de valores que maximiza o lucro no conjunto de treinamento. A primeira metade da partícula representa as operações envolvendo a posição comprado. “Distância para entrar na posição Comprado” representa o quão próxima a representação PAA da série original deve estar do padrão de entrada da posição comprado para provocar uma compra de ações e

Algoritmo 5: Regra de negócio para operar nas posições comprado e vendido.**Dados:**

D_Entrada_Comprado = Distância para entrar na posição comprado, definida pela melhor partícula

Dias_Venda = Limite da quantidade máxima de dias para operar na posição comprado, definido pela melhor partícula

D_Saida_Comprado = Distância para sair da posição comprado, definida pela melhor partícula

D_Entrada_Vendido = Distancia para entrar na posição vendido, definida pela melhor partícula

Dias_Compra = Limite da quantidade máxima de dias para operar na posição vendido, definido pela melhor partícula

D_Saida_Vendido = Distancia para sair da posição vendido, definida pela melhor partícula

P_Entrada_Comprado = representação PAA definida pela melhor partícula

P_Saída_Comprado = representação PAA definida pela melhor partícula

P_Entrada_Vendido = representação PAA definido pela melhor partícula

P_Saída_Vendido = representação PAA definido pela melhor partícula

Janela = janela da série temporal original retornada pela janela deslizante

```

1  início
2  se (Não entrou em posição COMPRADO e não entrou em posição VENDIDO) então
3  se ((MINDIST (P_Entrada_Comprado, PAA(Janela)) < D_Entrada_Comprado) E
4  (MINDIST (P_Entrada_Vendido, PAA(Janela)) < D_Entrada_Vendido)) então
5  se ((MINDIST (P_Entrada_Comprado, PAA(Janela)) < MINDIST
6  (P_Entrada_Vendido, PAA(Janela)) então
7  | Entra em posição COMPRADO e gera ordem de COMPRA
8  senão
9  | Entra em posição VENDIDO e gera ordem de VENDA
10 fim
11 senão
12 se (MINDIST (P_Entrada_Comprado, PAA(Janela)) < D_Entrada_Comprado)
13 então
14 | Entra em posição COMPRADO e gera ordem de COMPRA
15 se (MINDIST (P_Entrada_Vendido, PAA(Janela)) < D_Entrada_Vendido)
16 então
17 | Entra em posição VENDIDO e gera ordem de VENDA
18 fim
19 se (Entrou em posição COMPRADO) então
20 se (MINDIST (P_Saída_Comprado, PAA(Janela)) < D_Saída_Comprado) então
21 | Sai da posição COMPRADO e gera ordem de VENDA
22 se (Tempo dentro da posição COMPRADO > Dias_Venda) então
23 | Sai da posição COMPRADO e gera ordem de VENDA
24 se (Entrou em posição VENDIDO) então
25 se (MINDIST (P_Saída_Vendido, PAA(Janela)) < D_Saída_Vendido) então
26 | Sai da posição VENDIDO gera ordem de COMPRA
27 se (Tempo dentro da posição VENDIDO > Dias_Compra) então
28 | Sai da posição VENDIDO e gera ordem de COMPRA
29 fim

```

operar nessa posição. “Dias para Venda” representa o número máximo de dias em que o sistema se manterá com as ações na posição comprado. A melhor representação PAA do padrão de entrada na posição comprado encontrada pelo IDPSO, para usá-lo na etapa de teste. “Distância para sair da posição Comprado” o quão perto a representação PAA da série original deve estar do padrão de saída da posição comprado para provocar uma venda de ações, caso haja alguma compra nessa posição. Por fim, a melhor representação PAA encontrada pelo IDPSO para o padrão de saída da posição comprado.

Por sua vez, a segunda metade da partícula representa as operações envolvendo a posição vendido. “Distância para entrar na posição Vendido” representa o quão perto a representação PAA da série original deve estar do padrão de entrada em posição vendido para provocar uma venda de ações (aluguel de ações) e operar nessa posição. “Dias para Compra” representa o número máximo de dias em que o sistema se manterá com as ações alugadas. A melhor representação PAA do padrão de entrada em posição vendido encontrada pelo IDPSO, para usá-lo na etapa de teste. “Distância para sair da posição Vendido” o quão próxima a representação PAA da série original deve estar do padrão de saída da posição vendido para provocar uma compra de ações, caso haja alguma alugada nessa posição. Por fim, a melhor representação PAA encontrada pelo IDPSO para o padrão de saída da posição vendido.

Na fase de teste, o algoritmo calcula as similaridades entre cada um dos melhores padrões e a representação PAA da série original (subconjunto de teste). Então, essa distância é comparada com os parâmetros retornados pela fase de treinamento para decidir a operação e a estratégia a ser utilizada. A regra de decisão também calcula o número de dias em que o sistema está operando nas posições comprado ou vendido, para sair independentemente das distâncias quando se atinge o limite estabelecido nos “Dias para Venda” ou “Dias para Compra”, retornado pela fase de treinamento.

Um problema nesta nova abordagem está relacionado ao fato de que agora o algoritmo pode tomar duas decisões diferentes para investir, entrar na posição comprado e gerar uma ordem de comprar das ações, ou entrar na posição vendido, vendendo-as. Para lidar com o caso em que o método gera os dois sinais ao mesmo instante, o algoritmo opta em aplicar o sinal com base no menor valor da distância MINDIST entre a representação PAA da série temporal e cada um dos padrões de entrada da partícula, onde estão definidos os padrões de entrada da posição comprado e da posição vendido. Isto é, caso a MINDIST entre a representação PAA da série e o padrão de entrada em posição comprado seja menor que a distância calculada em relação ao padrão de entrada em posição vendido, a regra de negócio entra na posição comprado e gera uma ordem de compra, caso contrário entra em posição vendido e gera um sinal de venda.

3.4 PAA-IDPSO-CD

A técnica PAA-IDPSO-CD (Aproximação por Valor Agregado de Segmento (PAA) - Otimização por Enxame de Partículas Auto Adaptativa (IDPSO) com Detecção de Mudança de

Conceito (CD)), de forma diferente às demais versões apresentadas anteriormente, apresenta retreino do modelo baseado na detecção de mudança de conceito. Nos modelos apresentados anteriormente (seções 3.2 e 3.3), o treinado era efetuado apenas uma vez e as configurações retornadas pela melhor partícula são utilizadas no período de teste (uso do sistema). No entanto, isso representa uma limitação para o modelo, pois, como apresentado anteriormente, séries temporais se comportam como fluxos de dados que podem apresentar mudanças na distribuição dos dados e em seus padrões, ao longo do tempo (GAMA et al., 2014).

Por isso, nesta dissertação é proposta uma métrica para detecção de mudança de conceito que é utilizada para realizar retreino do modelo durante o período de teste. Espera-se que o retreino do modelo possa descobrir novos padrões representativos dos novos conceitos das séries temporais, a partir de seus dados mais recentes e, com isso, possa buscar obter melhores resultados quando comparados aos modelos que não utilizam o retreinamento. A métrica para detecção de mudança de conceito proposta pode ser utilizada tanto na estrutura do modelo que opera apenas na posição comprado quanto no modelo que opera em comprado-vendido.

A opção pelo uso de detecção de mudança de conceito em detrimento do uso de um modelo de retreino online se dá pelos seguintes motivos: 1) o retreino online é muito mais custoso e 2) pode haver perda de informação quando o retreino ocorre a cada iteração.

3.4.1 Métrica de detecção de mudança de conceito proposta

O presente estudo propõe uma métrica para a detecção de mudança de conceito. A motivação para usar essa opção é que as técnicas para a detecção de mudança apresentadas precisam de um classificador base para identificar a mudança de conceito. Por sua vez, o modelo proposto não utiliza classificadores, ele utiliza apenas padrões e regras de decisão, logo foi necessária uma adequação da ideia de mudança de conceito para o presente problema.

Para a detecção de mudança de conceito, o presente estudo utilizou a métrica *Drawdown*. Esta é definida como a queda relativa no preço de um máximo local para o próximo mínimo local, ignorando as auterações de preços entre os dois valores extremos. Em mercado financeiro, este valor é utilizado como uma variável que mede o risco. *Drawdown* capta, assim, a forma como quedas sucessivas podem influenciar umas às outras e construir, desta forma, um processo persistente ou quase persistente (JOHANSEN, 2003). Assim, no PAA-IDPSO-CD, o *drawdown* sobre o saldo obtido pode ser entendido como o erro do modelo, pois a queda no saldo ocorre quando o modelo não dispara uma ordem de compra ou de venda no melhor momento. O uso da métrica proposta apresenta como vantagem a definição dinâmica de seu valor, obtido a partir de cada treinamento/retreinamento para cada ação. Isso é fundamental, pois cada série temporal ou cada janela dentro da mesma série pode apresentar comportamento distinto.

De acordo com a teoria estatística, o erro diminui enquanto a distribuição dos dados for estacionária, aumentando quando ocorrer mudança na distribuição (GAMA et al., 2004). Logo, caso haja uma queda maior que a do treinamento, supõe-se que houve mudança no conceito e o

modelo deve ser retreinado a partir de uma janela de tempo com os valores mais recentes da série original, para melhor representar o novo conceito. Dessa forma, a técnica proposta funciona da seguinte: durante a fase de treinamento, calcula-se o *Máximo Drawdown* (representado por μ) do saldo resultante do modelo proposto, como pode ser observado na figura 3.6. Esse valor, por ser o maior erro obtido na etapa de treinamento, será utilizado como limiar para a realização do retreino na fase de teste, como pode ser visto na figura 3.7. Assim, caso, no teste, o saldo resultante do modelo caia a um valor percentual maior que μ , o modelo será retreinado. Então, a cada novo retreino serão calculados, a partir de uma nova otimização, tanto os novos valores dos parâmetros utilizados pelo PAA-IDPSO-CD quanto um novo valor de *Máximo Drawdown* (μ_i) a serem utilizados no teste.

Para não correr o risco de as janelas que são utilizadas para realizar o retreino terem tamanhos pequenos, o modelo trabalha com janelas de tempo de tamanho fixo nas suas etapas de treinamento e de retreinamento. Assim, durante a fase de teste cada novo valor decorrente da janela deslizante deve ser adicionado à janela de treinamento da seguinte forma: retira-se o valor mais antigo da janela e adiciona-se o novo valor. De tal forma que, quando for atingido o limite de μ_i no teste e for disparado um sinal de retreino, a janela de tempo utilizada para o retreino terá o mesmo tamanho da janela de treinamento inicial, porém, com os valores mais recentes da série original. Para o cálculo de μ_i , utiliza-se a Equação 4.1:

$$\mu_i = \frac{\text{SaldoMaximo}_i - \text{SaldoAtual}_i}{\text{SaldoMaximo}_i} \quad (\text{Eq.4.1})$$

No cálculo de (μ_i), utilizam-se apenas os saldos quando o modelo está fora do mercado. Isto é feito porque, admitindo que o sistema compre o máximo de ações com o seu saldo atual, quando efetuar a operação de compra, o saldo vai praticamente zerar. Logo esse saldo quase nulo, não pode ser utilizado para calcular o *drawdown*. Por exemplo, caso o investidor tenha R\$100.000,00 e deseje entrar em posição comprado com a cotação da ação esteja em R\$35,00, após a operação de compra ele irá ficar com 2857 ações e um saldo em dinheiro R\$5,00. Devido ao fato de o investidor ter uma quantidade de ações diferente de zero, este saldo em dinheiro de R\$5,00 não pode ser utilizado para o cálculo de μ_i , pois não representa o real saldo do investidor. O real saldo do investidor será calculado quando ocorrer a operação de venda das ações, isto é, quando estiver fora do mercado. Da mesma forma, quando entra na posição de vendido, o saldo irá subir a um valor acima do que realmente ele representa, este apenas será representado no posterior momento da compra. Por exemplo, se o investidor tiver R\$100.000,00 e entrar em posição vendido com 1000 ações a um preço de R\$35,00, após a operação de venda ficará com 1000 ações alugadas e um saldo de R\$135.000,00. Devido ao fato de o investidor ter uma quantidade de ações alugadas diferente de zero, este saldo em dinheiro de R\$135.000,00 não pode ser utilizado para o cálculo de μ_i , pois não representa o real saldo do investidor. O real saldo do investidor será calculado quando ocorrer a operação de compra das ações alugadas, isto é, quando estiver fora do mercado.

A Figura 3.6 apresenta um exemplo do cálculo de μ_i para o treinamento e retreinamentos, decorrente de uma execução para a ação AA com estratégia de investimento operando apenas em comprado. Nela, são apresentadas quedas menores do saldo do período na cor azul e a *Máximo Drawdown* μ_i na cor vermelha.

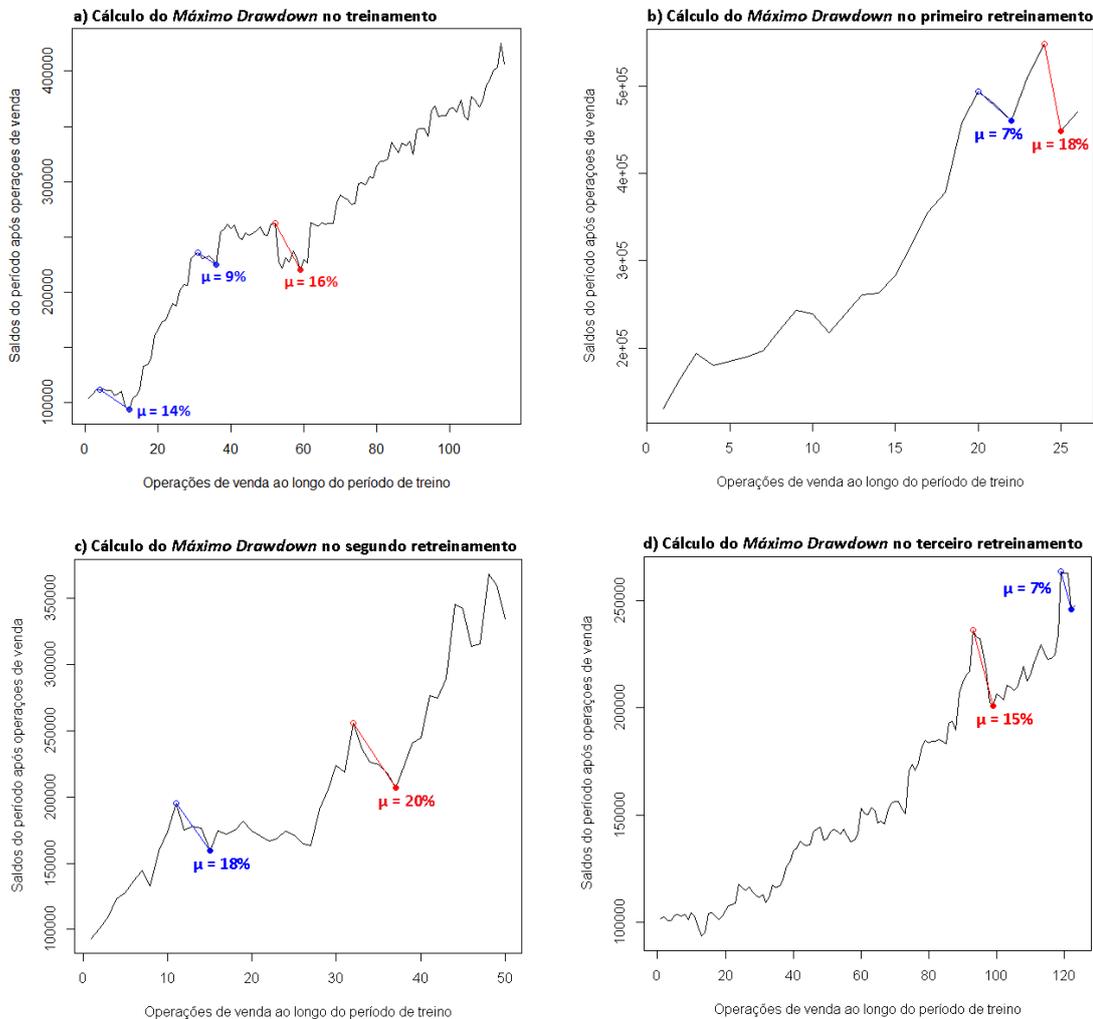


Figura 3.6: Cálculo da métrica de detecção de mudança de conceito proposta μ , nos conjuntos de treinamento e de retreinamentos. Em vermelho, destacado o *Máximo Drawdown*.

Por sua vez, a Figura 3.7 apresenta um exemplo do uso da detecção de mudança de conceito num conjunto de teste, ou seja, durante a operação do sistema após o treinamento, também decorrente da ação AA com estratégia de investimento operando apenas em comprado. Na figura, a tabela apresenta o valor dos *drawdowns* μ_i calculados, inicialmente a partir do conjunto de treino e, posteriormente, dos conjuntos de retreino. As linhas vermelhas apresentam os *drawdowns* do saldo decorrentes do conjunto de teste que resultaram em retreino do modelo, o momento em que ocorrem os retreinos estão destacados pelas setas azuis. O primeiro retreino ocorreu devido ao saldo do modelo no conjunto de teste ter atingido uma queda de 20%, maior que o $\mu_0 = 16\%$ derivado do conjunto de treinamento (Figura 3.6 a)). Na sequência, o segundo

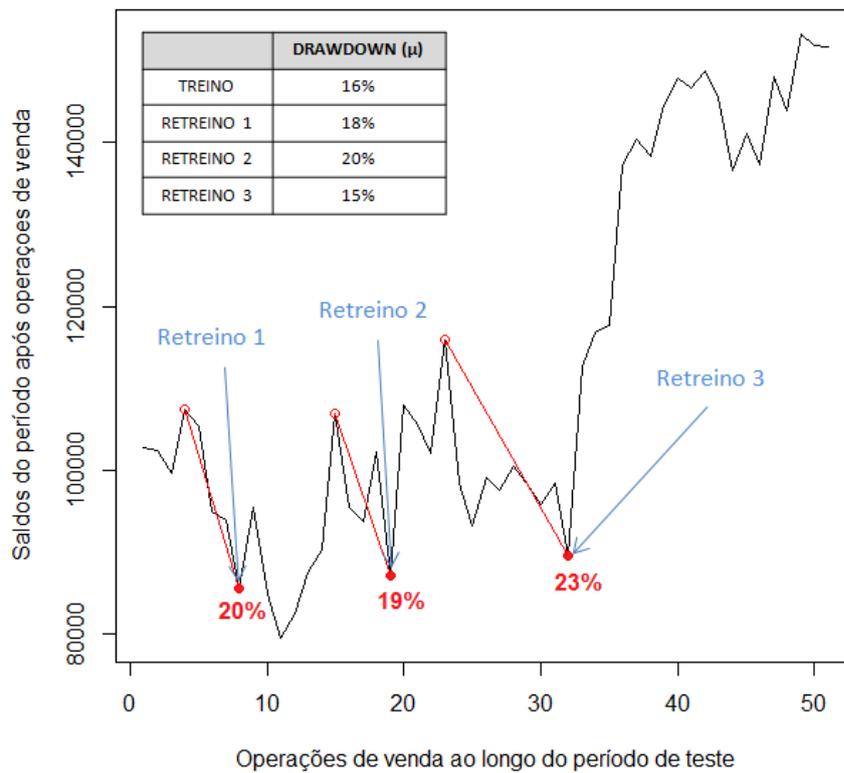


Figura 3.7: Detecção de mudança de conceito no conjunto de teste.

retreino ocorreu devido à queda de 19%, maior que o *Máximo drawdown* decorrente do primeiro retreino, que foi de $\mu_1 = 18\%$ (Figura 3.6 b)), e, assim, sucessivamente.

3.4.2 Método para redução da variância

Com o objetivo de reduzir a variância dos resultados obtidos pelo modelo, o presente estudo propõe o uso da partição do saldo inicial em modelos diferentes para a realização dos investimentos, buscando diminuir o risco das operações através da diversificação de modelos. Para isso, o saldo inicial SI é particionado em n modelos com configurações paramétricas diferentes da abordagem do PAA-IDPSO-CD utilizada. Ao contrário do que vinha sendo aplicado até agora, em que, a cada execução, o modelo começava com o saldo inicial, SI e eram aplicadas as regras de decisão, aqui, cada execução do PAA-IDPSO-CD corresponde a n configurações paramétricas diferentes do PAA-IDPSO-CD que começam com o saldo inicial de SI_x . O particionamento do saldo ocorre através da equação 4.2.

$$SI_x = \frac{SI}{n} \quad (Eq.4.2)$$

Assim, o desempenho ruim de um dos modelo pode ser compensado por um bom desempenho de outros modelos, na mesma execução. Com isso, os resultados obtidos não atingirão valores tão extremos quantos os apresentados pela abordagem em que não é utilizado o particionamento do saldo inicial, ou seja, apresentarão menor variância. Ao final de cada

execução, os saldos obtidos por cada modelo serão somados, resultando no saldo final da iteração SF , como o apresentado na equação 4.3.

$$SF = \sum_{x=1}^n S_x \quad (Eq.4.3)$$

A Figura 3.8 mostra o gráfico que apresenta a ideia do particionamento do saldo. Para cada execução do PAA-IDPSO-CD com aplicação do método de controle de variância proposto, o saldo inicial SI será particionado em n . Cada um desses saldos particionados irão rodar sobre um modelo PAA-IDPSO-CD diferente. Ao final, o somatório dos saldos resultantes de cada desses modelos será o resultado da iteração do PAA-IDPSO-CD.

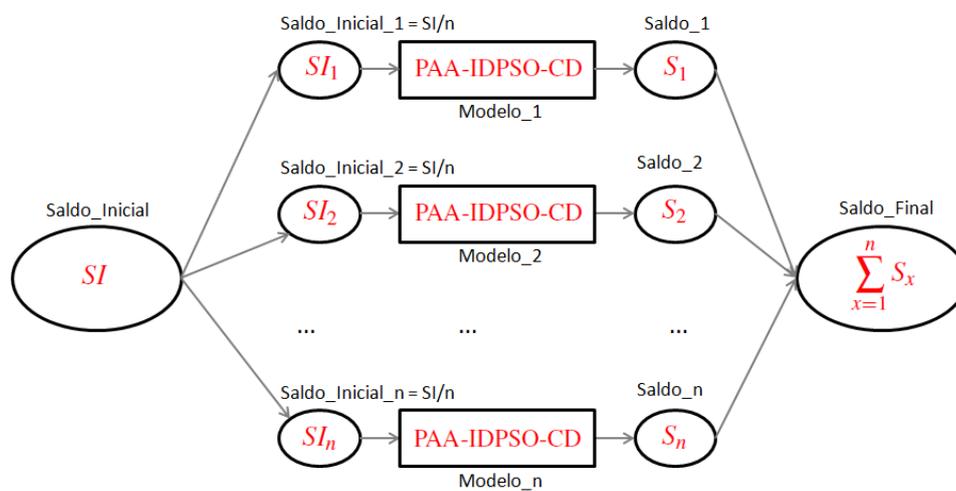


Figura 3.8: Particionamento do saldo inicial para redução de variância do modelo.

3.5 Considerações finais

Nesse capítulo, foi apresentada a evolução do modelo proposto até a versão final. Inicialmente, o modelo operava apenas em posições comprado e não havia retreino durante a fase de teste. Posteriormente, a nova regra de negócio apresentada operava nas posições comprado e vendido, mas sem retreino. Por fim, foram apresentadas uma métrica para detecção de mudança de conceito e um método para redução da variância do modelo. Enquanto este ocorre através do particionamento do saldo inicial e sua execução em modelos diferentes, aquele é utilizado para o retreino do modelo durante a fase de teste e que se adequa tanto à estrutura que opera apenas em posição comprado quanto na que opera nas duas estratégias.

4

Experimentos

4.1 Introdução

Neste capítulo são apresentados os experimentos e a discussão dos resultados obtidos. Os resultados foram comparados em termos de resultado financeiro final do período e de tempo de execução dos modelos. Um detalhe importante é o fato de o modelo proposto operar em escala diária. Isso acarreta uma menor influência das perdas de informações, decorrentes da redução de dimensionalidade resultante do PAA, se comparado ao universo de operações realizadas em alta frequência (cotação negócio-a-negócio, também conhecida como *tick-by-tick*), em que os lucros nas operações ocorrem devido à volatilidade no preço das ações (MURPHY, 1999). No entanto, não foi feito nenhum estudo comparativo sobre as perdas de informação decorrente da redução da dimensionalidade resultante do PAA

A análise é feita ação à ação, isto é, todo o saldo é aplicado ao modelo que opera apenas sobre uma ação, não sendo utilizada a ideia de portfólio de ações. Inicialmente, buscou-se demonstrar a viabilidade do modelo PAA-PSO como sendo um modelo mais eficaz que SAX-GA. Para isso, os resultados obtidos por ambas as abordagens são comparadas utilizando os mesmos valores de tamanho de janela (n) e o número de segmentos (k). Nessa comparação, usou-se a versão mais simples do PAA-PSO, com regra de negócio operando apenas em posição comprado e sem retreino.

Posteriormente, são apresentados os resultados comparativos entre todas as versões do modelo proposto até a sua versão proposta, o PAA-IDPSO-CD. Os resultados obtidos ainda são comparados com a estratégia *Buy and Hold* que é amplamente utilizado como uma referência no estudo de séries temporais financeiras aplicadas ao mercado financeiro e consiste em comprar as ações, mantendo-as na carteira de investimentos por um longo período de tempo (TEIXEIRA; OLIVEIRA, 2010), (BRASILEIRO et al., 2013), (CANELAS; NEVES; HORTA, 2012), (CANELAS; NEVES; HORTA, 2013).

São apresentados também uma análise e discussão sobre os resultados. Além disso, são aplicados testes estatísticos para medir se há diferença significativa entre os resultados obtidos. Por fim, são apresentadas as considerações finais.

4.2 Experimentos

As ações utilizadas foram as seguintes:

- Ações do índice no índice S&P500, Estados Unidos:
 - AA: Alcoa Inc.: empresa que trabalha com alumínio
 - AAPL: Apple Inc.: empresa da área de informática.
 - IBM: International Business Machines Corporation: empresa da área de informática.
 - KO: The Coca-Cola Company: empresa fabricante e comerciante de bebidas não-alcoólicas e concentrados de xaropes.
 - MCD: McDonald's Corp.: empresa responsável por uma rede internacional de restaurantes de fast food
 - MMM: 3M Company: empresa de tecnologia diversificada que trabalha em diversos setores, como Indústria, Transporte, Saúde, Consumo, Segurança entre outros
 - NKE: Nike Inc.: empresa de calçados, equipamentos esportivos, roupas, e acessórios
 - SBUX: Starbucks Corporation: cadeia de cafeterias
 - T: AT&T Inc.: empresa de telecomunicação
 - TIF: Tiffany & Co.: comércio de jóias
- Ações do índice no índice BM&FBOVESPA, Brasil:
 - ABEV3: AMBEV S.A.: empresa produção de bens de consumo
 - BBDC3: BANCO BRADESCO S.A.: banco
 - CRUZ3: SOUZA CRUZ S.A.: empresa produtora de cigarros
 - ELET3: ELETROBRAS S.A.: empresa do setor energético
 - EMBR3: EMBRAER S.A.: empresa fabricante de aviões
 - ITSA4: ITAU S.A.: banco
 - LIGT3: LIGHT S.A.: empresa do setor energético
 - PETR3: PETROBRAS S.A.: empresa do segmento de energia, com foco nas áreas de exploração, produção, refino, comercialização e transporte de petróleo, gás natural e seus derivados.
 - POMO4: MARCOPOLLO S.A.: empresa fabricante de carrocerias de ônibus
 - USIM5: USIMINAS S.A.: empresa do setor siderúrgico

Essas ações são de diferentes setores econômicos, apresentarem comportamento diferente e foram escolhidas ao acaso porque, principalmente, no caso das ações do BM&FBOVESPA, não haviam dados das cotações do período utilizado para todas as ações do mercado. Os dados das cotações das ações foram obtidos através do Yahoo Finance, tanto em relação ao índice S&P500¹ quanto ao índice BM&FBOVESPA². As operações ocorreram em escala diária e utilizou-se a cotação ajustada de fechamento como preço das ações. O período utilizado foi janeiro de 2000 a junho de 2010, o comportamento das séries temporais durante esse período pode ser observado nas figuras A.1 a A.7, presentes no Apendice A. De janeiro de 2000 a dezembro de 2004, período de treinamento, e de janeiro de 2005 a junho de 2010, o período de teste. Os experimentos foram executados 50 vezes e, para cada execução, o saldo inicial em dinheiro foi de \$ 100.000,00 (cem mil reais no caso das ações do mercado brasileiro e cem mil dólares para as ações do mercado norteamericano). Além disso, com o objetivo de obter uma melhor simulação do mercado de ações real, foi utilizado o valor de \$9,80 como custo de transação³ (valor cobrado tanto nas operações de compra quanto de venda).

Em relação ao algoritmo PAA-PSO utilizando regra de decisão operando apenas na estratégia comprado, os parâmetros utilizados são os seguintes:

- “Distância para Compra”: valores do tipo ponto flutuante no intervalo [0, 40];
- “Distância para Venda”: valores do tipo ponto flutuante no intervalo [0, 40];
- “Dias para Venda”: valores inteiros no intervalo [1, 50];
 - No decorrer da otimização esse valor é trabalhado como ponto flutuante, para uma melhor adequação ao algoritmo de otimização utilizado. A conversão para valor inteiro ocorre na hora de esse valor ser utilizado pela regra de decisão do modelo.
- Para cada valor \bar{x}_1 do padrão $PAA = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k]$ a ser descoberto: valores do tipo ponto flutuante no intervalo [-4.5, 4.5], devido à sua padronização.

Em relação à regra de negócio que opera nas estratégias comprado de vendido, os parâmetros utilizados são os seguintes:

- “Distância para entrar na posição Comprado”: valores do tipo ponto flutuante no intervalo [0, 40];
- “Dias para Venda”: valores inteiros no intervalo [1, 50];
 - No decorrer da otimização esse valor é trabalhado como ponto flutuante, para uma melhor adequação ao algoritmo de otimização utilizado. A conversão para valor inteiro ocorre na hora de esse valor ser utilizado pela regra de decisão do modelo.

¹<http://finance.yahoo.com/>

²<https://br.financas.yahoo.com/>

³<https://www.rico.com.vc/servicos/custos-de-operacao>

- Para cada valor x_i do padrão de entrada em posição Comprado $PAA = [x_1^{EL}, x_k^{EL}, \dots, x_k^{EL}]$: valores do tipo ponto flutuante no intervalo $[-4.5, 4.5]$, devido a padronização no seu valor.
- “Distância para sair da posição Comprado”: valores do tipo ponto flutuante no intervalo $[0, 40]$;
- Para cada valor x_i do padrão de saída para a estratégia Comprado $PAA = [x_1^{SL}, x_k^{SL}, \dots, x_k^{SL}]$: valores do tipo ponto flutuante no intervalo $[-4.5, 4.5]$, devido a padronização no seu valor.
- “Distância para entrar na posição Vendido”: valores do tipo ponto flutuante no intervalo $[0, 40]$;
- “Dias para Compra”: valores inteiros no intervalo $[1, 50]$;
 - No decorrer da otimização esse valor é trabalhado como ponto flutuante, para uma melhor adequação ao algoritmo de otimização utilizado. A conversão para valor inteiro ocorre na hora de esse valor ser utilizado pela regra de decisão do modelo.
- Para cada valor x_i do padrão de entrada na posição Vendido $PAA = [x_1^{ES}, x_k^{ES}, \dots, x_k^{ES}]$: valores do tipo ponto flutuante no intervalo $[-4.5, 4.5]$, devido a padronização no seu valor.
- “Distância para sair da posição Vendido”: valores do tipo ponto flutuante no intervalo $[0, 40]$;
- Para cada valor x_i do padrão de saída para a estratégia Vendido $PAA = [x_1^{SS}, x_k^{SS}, \dots, x_k^{SS}]$: valores do tipo ponto flutuante no intervalo $[-4.5, 4.5]$, devido a padronização no seu valor.

Para o PSO foram usados os parâmetros padrões: velocidade máxima = 1.0, $c_1 = c_2 = 2.00$, inércia = 0.9 (HASSAN et al., 2005). O número de partículas no enxame = 50, número máximo de iterações = 1000, critério de parada de 100 iterações sem melhoria no fitness. Para a definição do número de partículas no enxame foram efetuados testes com 20, 50 e 100 partículas. Como pode ser observado na Figura 4.1, o fitness obtido pelo algoritmo com 50 partículas melhor que o fitness obtido pela execução com 20 partículas é semelhante ao fitness do algoritmo com 100 partículas (executado na regra de negócio operando apenas comprado para a ação AA). Assim, usar 50 partículas foi capaz de chegar a um bom resultado, sem necessidade de extrapolar a quantidade de partículas, o que tornaria o algoritmo mais custoso.

Os mesmos valores foram utilizados para GA, além de taxa de mutação = 10% (CANELAS; NEVES; HORTA, 2012). Como a taxa de cruzamento não foi informada em CANELAS; NEVES; HORTA (2012), o presente estudo utilizou um valor de 50%, como apresentado por HASSAN et al. (2005). Este estudo utilizou GA com 50 indivíduos para fazer uma comparação justa com os 50 partículas de PSO. Além disso, durante a execução do experimentos, observou-se que a GA com 50 indivíduos apresentou um bom equilíbrio entre resultados e custo computacional quando comparado ao GA com 500 indivíduos, como o usado por CANELAS; NEVES;

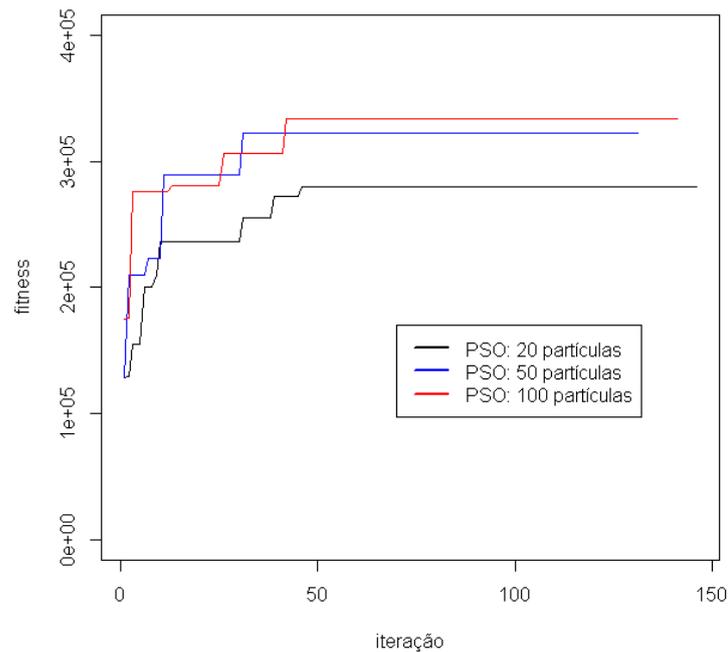


Figura 4.1: Convergência PSO com 20, 50 e 100 partículas.

HORTA (2012). Isso pode ser observado na Figura 4.2, que mostra a convergência da execução do algoritmo para a regra de negócio operando apenas em posições comprado, para a ação AA.

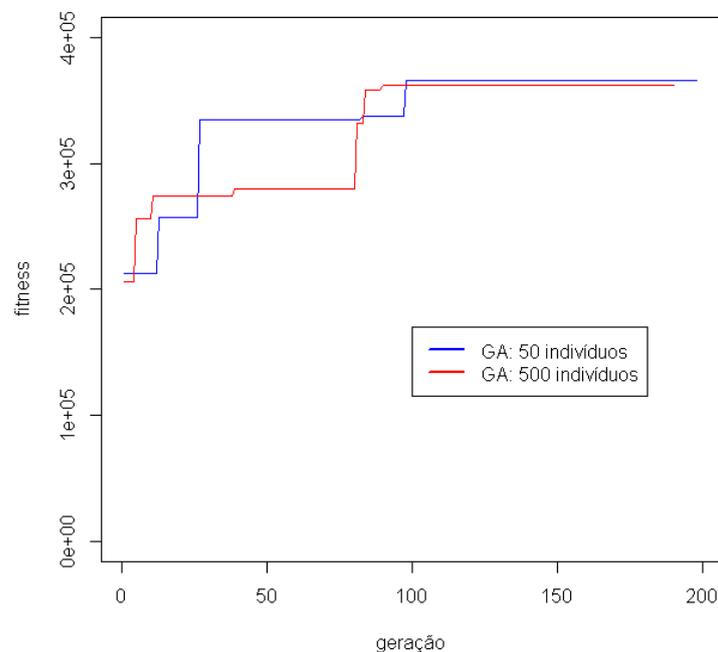


Figura 4.2: Convergência Algoritmo Genético com 50 e 500 indivíduos.

Para o IDPSO, também foram utilizados o número de partículas no enxame = 50, número máximo de iterações = 500, critério de parada de 100 iterações sem melhoria no fitness. Assim como no modelo original, $w_{inicial} = 0.9$ e $w_{final} = 0.4$ (ZHANG; XIONG; ZHANG, 2013). O número máximo de iterações foi reduzido para 500, pois foi observado que a convergência

ocorria por volta de 273 iterações, com desvio padrão de 138. Com isso, ao realizar essa redução tem-se um aumento no poder de busca no espaço de soluções, por conta do valor da função de detecção $\varphi(t)$.

Com o objetivo de obter um estudo comparativo, foram realizados Testes t Pareado com nível de confiança de 95% (MONTGOMERY; RUNGER, 2012). O Teste t é utilizado quando os dados da estatística de teste segue uma distribuição normal, mas a variância da população σ^2 é desconhecida. A opção por esse teste decorre da Teorema do Limite Central. De acordo com esse teorema, a medida que o tamanho n da amostra aumenta, a sua distribuição amostral tende a apresentar uma distribuição normal, em geral, $n = 30$. Como o presente estudo apresenta uma amostra de tamanho 50, pode-se aplicar o Teorema do Limite Central e admitir que a distribuição amostral tende à distribuição normal, com isso, pode-se utilizar o Teste t . A opção por usá-lo na sua forma pareada decorre do fato de as amostras de pares dos modelos serem tomadas sob condições homogêneas (MONTGOMERY; RUNGER, 2012).

Um teste de hipótese foi realizado para verificar se existe diferença significativa entre os resultados obtidos pelos modelos:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

Em que H_0 é a hipótese nula, e H_1 é a hipótese alternativa, μ_1 é a média da amostra A_1 e μ_2 é a média da amostra A_2 , com 95% de nível de confiança. Para o teste estatístico, caso $p\text{-valor} < 0,05$, rejeita-se H_0 , isto é, existe diferença significativa e A_1 tem melhores resultados que A_2 . Caso contrário, não se rejeita H_0 , ou seja, não existe diferença significativa entre A_1 e A_2 .

4.2.1 Comparação PAA-PSO com SAX-GA ambas com regra de negócio operando apenas em posição comprado

O objetivo, aqui, é comparar o PAA-PSO com o SAX-GA tanto em relação ao resultado financeiro final do período, como em tempo de execução, utilizando os mesmos valores de tamanho de janela (n) e o número de segmentos (k). Em relação ao tamanho do alfabeto α utilizado pelo SAX, mudanças foram feitas em seu valor ($\alpha = 6$; $\alpha = 12$; $\alpha = 18$), com a finalidade de demonstrar a importância deste parâmetro para o SAX. A idéia, com isso, foi demonstrar uma das vantagens do modelo proposto, uma vez que o PAA não utiliza este parâmetro. Os parâmetros do modelo (tamanho da janela n e quantidade de segmentos k) para cada ação são escolhidos pelo usuário, no presente estudo, os valores foram escolhidos de forma empírica. Para esses experimentos, foram utilizadas apenas as seguintes ações, com suas respectivas configurações:

- AA: $n = 90$ e $k = 9$

- ABEV3: $n = 100$ e $k = 10$
- ITSA4: $n = 150$ e $k = 15$
- USIM5: $n = 90$ e $k = 9$

A tabela 4.1 apresenta um comparativo entre os resultados obtidos em 50 execuções do PAA-PSO e do SAX-GA, em suas três variações α . As colunas média e desvio padrão estão relacionados ao resultado financeiro final do período e a coluna tempo de execução apresenta o o tempo de execução para as 50 execuções, em segundos. Na tabela, para cada ação, o modelo que apresentou a melhor média está destacado.

Tabela 4.1: Comparativo dos resultados obtidos pelo PAA-PSO e SAX-GA nas 50 execuções.

Ação: AA	MÉDIA	DESVIO PADRÃO	TEMPO DE EXECUÇÃO
PAA-PSO	99768.78	63685.57	677 s
SAX-GA ($\alpha = 6$)	69050.81	40932.09	1024 s
SAX-GA ($\alpha = 12$)	79546.56	56741.32	1163 s
SAX-GA ($\alpha = 18$)	74218.65	51323.33	1800 s
Ação: ABEV3	MÉDIA	DESVIO PADRÃO	TEMPO DE EXECUÇÃO
PAA-PSO	250276.4	271365.3	708 s
SAX-GA ($\alpha = 6$)	156373.9	94288.54	791 s
SAX-GA ($\alpha = 12$)	220587.4	185713.6	1012 s
SAX-GA ($\alpha = 18$)	224884.7	221718.4	1098 s
Ação: ITSA4	MÉDIA	DESVIO PADRÃO	TEMPO DE EXECUÇÃO
PAA-PSO	201524.2	83778.99	582 s
SAX-GA ($\alpha = 6$)	174428.8	72037.42	1067 s
SAX-GA ($\alpha = 12$)	162349.7	86867.51	1503 s
SAX-GA ($\alpha = 18$)	159883.6	75296.98	1591 s
Ação: USIM5	MÉDIA	DESVIO PADRÃO	TEMPO DE EXECUÇÃO
PAA-PSO	299273.6	144351.9	632 s
SAX-GA ($\alpha = 6$)	242050	125176.5	720 s
SAX-GA ($\alpha = 12$)	274583.4	157360	925 s
SAX-GA ($\alpha = 18$)	222664.8	102977.3	1024 s

Em termos de custo computacional, a tabela 4.1 mostrou que PAA-PSO apresentou melhor desempenho (menor tempo de execução) do que todas as versões do SAX-GA, e em todas as ações. Essa é outra vantagem para o modelo proposto, já que os investidores buscam modelos que respondam de forma rápida às mudanças do mercado.

O modelo proposto apresentou melhor desempenho por duas razões: em primeiro lugar, o treinamento utilizando o PSO foi mais rápido do que o GA. Isso era de se esperar, pois uma das premissas pela opção do uso do PSO foi apresentar boa eficiência computacional, quando utilizado para resolver problemas com variáveis contínuas (HASSAN et al., 2005). Em segundo lugar, o PAA não precisa de pesquisa na tabela de pontos de interrupção para o cálculo de

distâncias. Além disso, com o aumento do tamanho do alfabeto, o algoritmo SAX-GA torna-se mais custoso devido ao maior número de operações decorrentes do aumento do tamanho da Tabela Ponto de Interrupção.

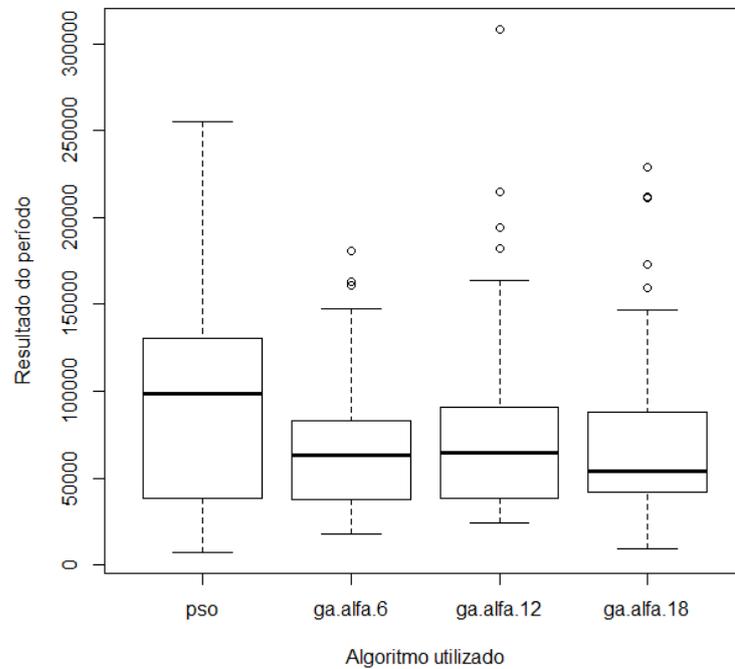


Figura 4.3: *Boxplot* do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação AA.

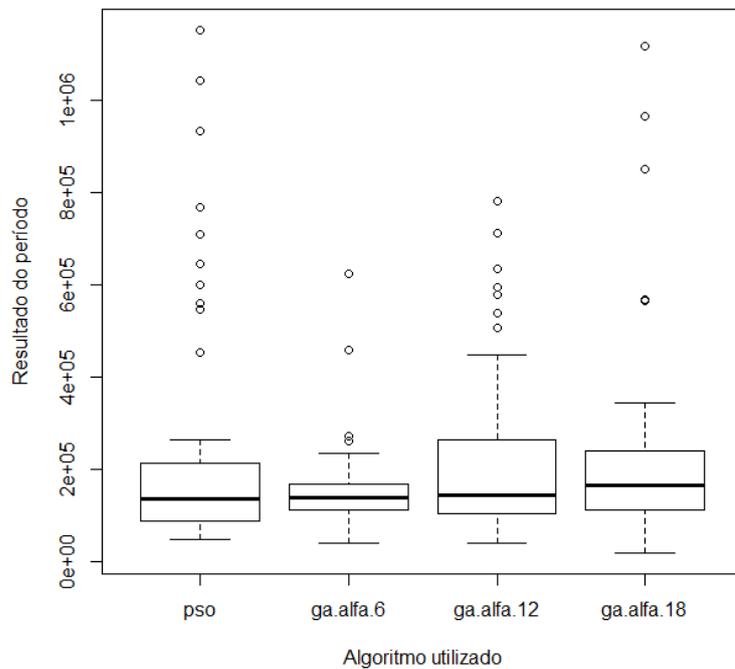


Figura 4.4: *Boxplot* do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação ABEV3.

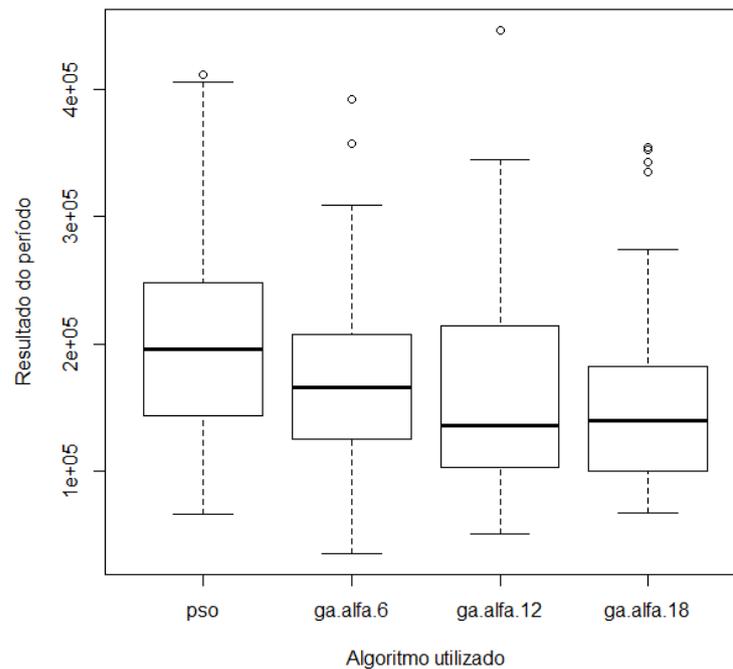


Figura 4.5: *Boxplot* do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação ITSA4.

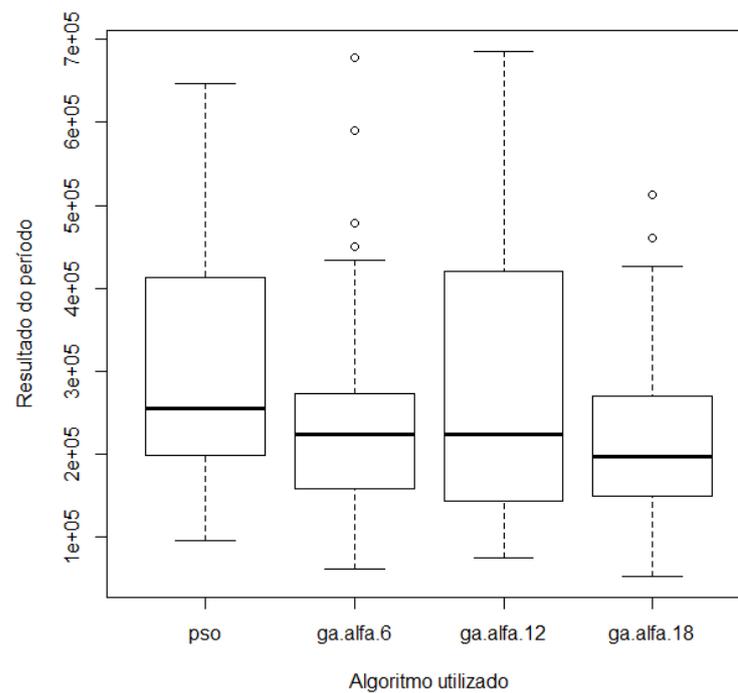


Figura 4.6: *Boxplot* do resultado financeiro final do período obtido pelos modelos PAA-PSO e SAX-GA nas 50 execuções para a ação USIM5.

Quanto ao estudo estatístico comparativo, como está sendo utilizado o Teste t Pareado com confiança de 95%, tem-se PAA-PSO como sendo A_1 e cada uma das versões do SAX-GA como A_2 . A tabela 4.2 mostra os resultados estatísticos de quando houve diferença significativa ou não entre os modelos. Para uma melhor visualização dos resultados obtidos pelo modelos, as

figuras 4.3, 4.4, 4.5, 4.6 apresentam, respectivamente, o *boxplot* para as 50 execuções relativas às ações AA, ABEV3, ITSA4 e USIM5.

Tabela 4.2: Testes estatísticos para o comparativo dos resultados obtidos pelo PAA-PSO e SAX-GA nas 50 execuções, com nível de confiança de 95%.

Ação: AA	RESULTADOS DOS TESTES ESTATÍSTICOS
SAX-GA ($\alpha = 6$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
SAX-GA ($\alpha = 12$)	Não houve diferença significativa entre o SAX-GA e PAA-PSO
SAX-GA ($\alpha = 18$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
Ação: ABEV3	RESULTADOS DOS TESTES ESTATÍSTICOS
SAX-GA ($\alpha = 6$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
SAX-GA ($\alpha = 12$)	Não houve diferença significativa entre o SAX-GA e PAA-PSO
SAX-GA ($\alpha = 18$)	Não houve diferença significativa entre o SAX-GA e PAA-PSO
Ação: ITSA4	RESULTADOS DOS TESTES ESTATÍSTICOS
SAX-GA ($\alpha = 6$)	Não houve diferença significativa entre o SAX-GA e PAA-PSO
SAX-GA ($\alpha = 12$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
SAX-GA ($\alpha = 18$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
Ação: USIM5	RESULTADOS DOS TESTES ESTATÍSTICOS
SAX-GA ($\alpha = 6$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)
SAX-GA ($\alpha = 12$)	Não houve diferença significativa entre o SAX-GA e PAA-PSO
SAX-GA ($\alpha = 18$)	PAA-PSO obteve resultados estatisticamente melhores quando comparados com SAX-GA ($p - valor < 0,05$)

A partir da tabela 4.2, na ação AA, pode ser observado que para $\alpha = 12$ não houve diferenças significativas entre os resultados do PAA-PSO e do SAX-GA. No entanto, para $\alpha = 6$ e $\alpha = 18$, a hipótese de igualdade entre os resultados foi rejeitada e, com isso, a abordagem PAA-PSO gera resultados significativamente melhores.

Para o ação ABEV3, para $\alpha = 6$, a hipótese de igualdade entre PAA-PSO e SAX-GA foi rejeitada, assim, a abordagem PAA-PSO gera resultados estatisticamente melhores. Por outro lado, para $\alpha = 12$ e $\alpha = 18$, não foram encontradas diferenças significativas entre os resultados dos modelos.

Em ITSA4, para $\alpha = 12$ e $\alpha = 18$, PAA-PSO apresentou resultados estatisticamente melhores do que SAX-GA. No entanto, para $\alpha = 6$, os resultados obtidos por PAA-PSO e SAX-GA não têm diferença significativa.

Finalmente, na ação USIM5, verificou-se que para $\alpha = 12$ não foram observadas diferenças significativas entre os resultados das abordagens PAA-PSO e SAX-GA. No entanto, para $\alpha = 6$ e $\alpha = 18$ a hipótese de igualdade entre os resultados foi rejeitada, com isso, a abordagem PAA-PSO obteve resultados estatisticamente melhores.

Com o intuito de facilitar o entendimento do sistema, a Figura 4.7 apresenta o padrão PAA representativo da série temporal AA descoberto e utilizado pela abordagem PAA-PSO, em uma de suas 50 execuções. Caso, durante o uso do sistema, apareça a representação PAA de uma janela de série temporal similar ao padrão da figura, isto é, que apresente uma distância menor que o parâmetro “Distância para Compra” retornado pelo método na etapa de treinamento, espera-se que o comportamento posterior seja o de subida nos preços. Dessa forma, é disparado uma ordem de compra para tentar obter lucro na operação de venda posterior a um preço maior que o de compra. Por sua vez, a Figura 4.8 apresenta o comportamento do saldo ao longo do uso do sistema a partir do padrão PAA apresentado. Como pode ser observado, a execução resultou num lucro das operações de aproximadamente 67% sobre o saldo inicial, uma vez que o saldo final do modelo foi \$166.941,32.

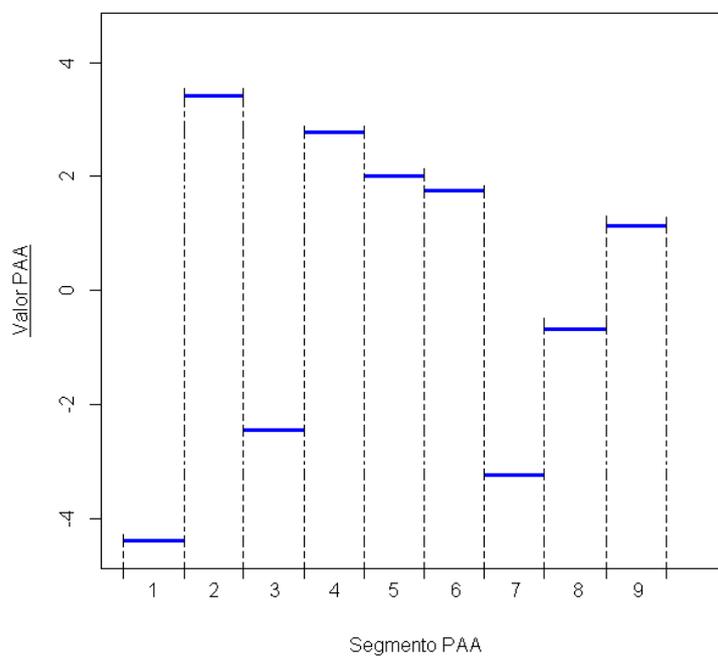


Figura 4.7: Padrão PAA representativo da série temporal AA, em uma execução, com tamanho da janela de série temporal (n) = 90 e quantidade de segmentos (k) = 9.

Nessa subseção foram comparadas as abordagens PAA-PSO e SAX-GA, ambas com regra de negócio operando apenas em posição Comprado. O PAA-PSO apresentou resultados iguais ou superiores ao SAX-GA, em termos de resultado financeiro final do período, a um custo computacional menor. Isso ratificou a viabilidade e a validade do modelo proposto.

Espera-se que no cenário de regra de negócio operando nas posições Comprado e Vendido esse comportamento se repita. Além disso, como a versão com retreino apresenta resultados muito superiores à versão sem retreino, não há mais a necessidade de estudos mais aprofundados dos resultados obtidos pelo SAX-GA e, com isso, esse modelo não será mais utilizado como base de comparação. Assim, a partir desse momento, apenas as versões do modelo proposto serão utilizados para efeito de comparação.

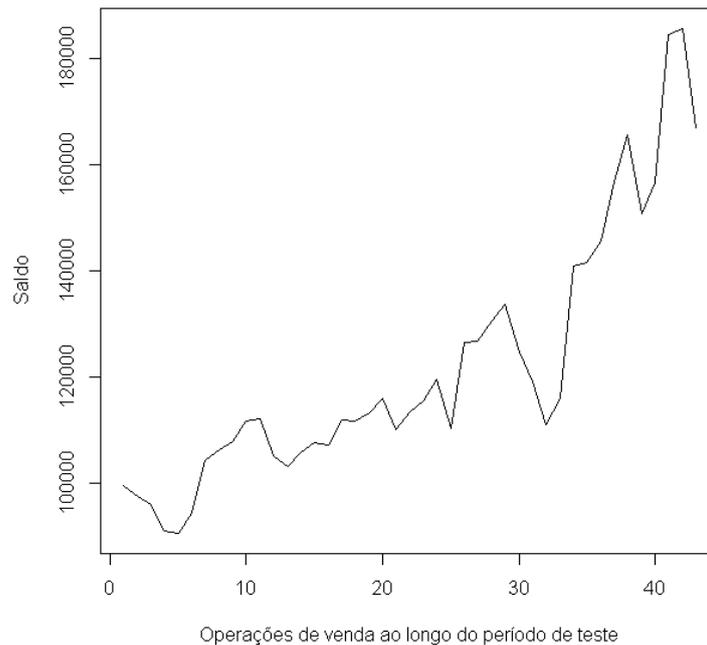


Figura 4.8: Saldo do modelo ao longo do uso do sistema a partir do padrão PAA apresentado.

4.2.2 Comparação PAA-IDPSO e PAA-PSO sem retreino

O objetivo, nesta subseção, é realizar a comparação entre o PAA-PSO e o PAA-IDPSO sem utilizar o retreino baseado na detecção de mudança de conceito, tanto na estratégia de operando apenas na posição comprado quanto naquela operando em comprado e vendido. As variáveis a serem comparadas mais uma vez são o resultado financeiro final do período e o tempo de execução, utilizando os mesmos valores de tamanho de janela (n) e o número de segmentos (k). Aqui, serão utilizadas todas as ações definidas anteriormente com o objetivo e obter a maior representatividade dos modelos. Mais uma vez, os valores dos parâmetros escolhidos de forma empírica e as configurações utilizadas para o PAA foram as seguintes:

- $n = 90$ e $k = 9$ para : AA, EMBR3, IBM, LIGT3, MCD, NKE, POMO4, T, TIF, USIM5
- $n = 100$ e $k = 10$ para: AAPL, ABEV3, BBDC3, CRUZ3, ELET3
- $n = 150$ e $k = 15$ para: ITSA4, KO, MMM, PETR3, SBUX

A tabela 4.3 apresenta um quadro comparativo entre os resultados obtidos pelas abordagens utilizando PAA-PSO e PAA-IDPSO, ambas operando nos dois tipos de estratégia e sem utilizar retreino. Na primeira coluna, PSO-COMP representa a abordagem PAA-PSO operando apenas na posição comprado, PSO-COMP-VEND, a abordagem PAA-PSO operando tanto em comprado quanto em vendido, IDPSO-COMP representa a abordagem PAA-IDPSO operando apenas na posição comprado, IDPSO-COMP-VEND, a abordagem PAA-IDPSO operando tanto em comprado quanto em vendido. A coluna média e desvio estão relacionadas aos resultados

finais do período obtidos nas 50 execuções e a última coluna, tempo, apresenta o tempo gasto, em segundos, para executá-las. Em negrito estão as abordagens que obtiveram os melhores resultados, para cada ação.

Tabela 4.3: Comparativo dos resultados obtidos pelo PAA-PSO e PAA-IDPSO, sem detecção de mudança de conceito, para as 50 execuções.

ABORDAGEM	AA			AAPL		
	MÉDIA	DESVIO	TEMPO	MÉDIA	DESVIO	TEMPO
PSO-COMP	99768.78	63685.57	677 s	332238.6	243801.2	683 s
PSO-COMP-VEND	170528.3	126457.9	2050 s	381395.1	204790.9	2021 s
IDPSO-COMP	115217.2	65146.13	698 s	387539.9	222543.1	747 s
IDPSO-COMP-VEND	154253.5	104579	2694 s	385800.4	211605.6	2518 s
ABORDAGEM	ABEV3			BBDC3		
PSO-COMP	250276.4	271365.3	708 s	202148.7	170700	704 s
PSO-COMP-VEND	277599.4	271087.8	2479 s	160392.1	123804.1	2189 s
IDPSO-COMP	266165.1	308320.3	797 s	190809.5	151992.8	724 s
IDPSO-COMP-VEND	261175.3	204496.5	2591 s	165796.8	119498.4	2391 s
ABORDAGEM	CRUZ3			ELET3		
PSO-COMP	451166.6	189570.3	698 s	160982.7	84406.16	648 s
PSO-COMP-VEND	458511.2	145911.1	2172 s	230973.4	95556.06	2139 s
IDPSO-COMP	497083.8	203576.1	788 s	196339.3	91578.67	725 s
IDPSO-COMP-VEND	508167	156059.2	2557 s	263682.1	121488.1	2562 s
ABORDAGEM	EMBR3			IBM		
PSO-COMP	58770.9	25356.15	622 s	90112.1	16555.27	600 s
PSO-COMP-VEND	46728.39	29417.58	2138 s	80452.84	27501.08	1965 s
IDPSO-COMP	62163.78	30115.18	756 s	94241.34	17640.43	769 s
IDPSO-COMP-VEND	55877.44	31781.45	2475 s	85878.75	21780.43	2518 s
ABORDAGEM	ITSA4			KO		
PSO-COMP	201524.2	83778.99	582 s	102098.8	24584.47	910 s
PSO-COMP-VEND	224960.7	87538.98	2905 s	104627.1	32802.16	2848 s
IDPSO-COMP	221829.1	82449.8	729 s	99840.97	28204.66	1030 s
IDPSO-COMP-VEND	211829.9	55620.4	3361 s	103100.9	35626.02	3543 s
ABORDAGEM	LIGT3			MCD		
PSO-COMP	79654.37	32401.7	630 s	160548.2	28056.61	657 s
PSO-COMP-VEND	86145.83	44578.42	2153 s	164585.2	35921.94	2117 s
IDPSO-COMP	92389.6	49417.18	647 s	164168.4	37975.97	775 s
IDPSO-COMP-VEND	84890.34	51183.52	2215 s	157087.1	41295.08	2647 s
ABORDAGEM	MMM			NKE		
PSO-COMP	109137.6	29194.61	1004 s	191732.6	47952.95	641 s
PSO-COMP-VEND	117677.9	30383	3212 s	199226	54977.9	2188 s
IDPSO-COMP	109811.7	27258.66	1032 s	174519.4	50158.37	675 s
IDPSO-COMP-VEND	122013.4	41373.45	3690 s	179803.1	45248.11	2383 s
ABORDAGEM	PETR3			POMO4		
PSO-COMP	407722.2	197995.8	917 s	302802.4	151613.8	609 s
PSO-COMP-VEND	358546.9	189662.2	2838 s	293930.3	156362.1	1852 s
IDPSO-COMP	335940.4	149961.5	1025 s	279987.7	158669.1	644 s
IDPSO-COMP-VEND	383617.7	179388.1	3479 s	336441	195065	2309 s

Tabela 4.3: Continuação.

Abordagem	SBUX			T		
PSO-COMP	114435.2	38738.79	853 s	118339.7	35420.37	641 s
PSO-COMP-VEND	128967.8	38602.49	3144 s	105365.2	28161.72	2060 s
IDPSO-COMP	114074.9	40900.14	1038 s	116505.1	31009.08	654 s
IDPSO-COMP-VEND	130642.1	34675.72	3459 s	105934.5	24768.37	2545 s
Abordagem	TIF			USIM5		
PSO-COMP	108767.6	50135.77	628 s	299273.6	144351.9	632 s
PSO-COMP-VEND	97522.11	76333.16	1887 s	308608.5	174813.4	2065 s
IDPSO-COMP	122929.7	53444.65	674 s	255600.3	122044.2	649 s
IDPSO-COMP-VEND	103144.2	42484.92	2076 s	276427.2	131467.2	2180 s

Em relação ao custo computacional, como pode ser observado, utilizar a estratégia comprado-vendido é muito mais custosa que operar apenas em comprado. Esse fato decorre, principalmente, da estrutura da partícula e das regras de decisão que ficam mais complexas na estratégia comprado-vendido, o que torna tanto o treinamento quanto o teste mais custosos quando comparada à operacionalização utilizando apenas a posição comprado. Além disso, o IDPSO é um pouco mais custoso quando comparado ao PSO, devido à atualização dos parâmetros ao longo das iterações no IDPSO.

Por sua vez, quanto aos resultados finais do período obtidos, a estatística comparativa foi executada da seguinte forma: primeiramente, fixando a abordagem utilizada (PAA-PSO ou PAA-IDPSO) e variando as estratégias de operacionalização, na sequência, fixando as estratégias de operacionalização e variando a abordagem utilizada. A tabela 4.4 apresenta os resultados para esses testes estatísticos.

Tabela 4.4: Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO e PAA-IDPSO, sem detecção de mudança de conceito, nas 50 execuções, com nível de confiança de 95%

ABORDAGENS	RESULTADOS DOS TESTES ESTATÍSTICOS
PSO-Comp X PSO-Comp-Vend	PSO-Comp: resultados estatisticamente melhores ($p - valor < 0,05$) em: IBM, T PSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, ELET3
IDPSO-Comp X IDPSO-Comp-Vend	IDPSO-Comp resultados estatisticamente melhores ($p - valor < 0,05$) em: IBM, T, TIF IDPSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, ELET3, SBUX
PSO-Comp X IDPSO-Comp	Não houve diferença significativa entre as abordagens em nenhuma ação
PSO-Comp-Vend X IDPSO-Comp-Vend	Não houve diferença significativa entre as abordagens em nenhuma ação

Como pode ser visto a partir da tabela 4.4, fixando-se no PSO e variando as regras de operacionalização, a estratégia apenas comprado obteve melhores resultados para as ações IBM e T; por sua vez, a estratégia comprado-vendido obteve melhores resultados nas ações AA

e ELET3. Por sua vez, ao fixar o IDPSO e variando as formas operacionalização, a estratégia comprado obteve melhores resultados para as ações IBM, T, TIF; por sua vez, a estratégia comprado-vendido obteve melhores resultados nas ações AA, ELET3 e SBUX. Por fim, nas comparações entre as abordagens PSO e IDPSO, não houve diferença estatisticamente significativa para nenhuma ação. Para esses resultados, não há evidências para se dizer se um método é melhor que o outro.

4.2.3 Comparação PAA-IDPSO-CD e PAA-PSO-CD com retreino baseado na detecção de mudança de conceito

O objetivo, nesta subseção, é realizar três comparações: a comparação dos resultados obtidos com aqueles apresentados pela abordagem *Buy and Hold (B&H)*; a comparação dos modelos que utilizaram retreino baseado na mudança de conceito com aqueles que não utilizaram; por fim, a comparação entre o PAA-PSO-CD e o PAA-IDPSO-CD tanto na estratégia operando apenas na posição comprado quanto na comprado e vendido. As variáveis a serem comparadas são, novamente, o resultado financeiro final do período e o tempo de execução, utilizando os mesmos valores de tamanho de janela (n) e o número de segmentos (k). Aqui, serão utilizados todas as ações definidas anteriormente com o objetivo de obter resultados mais amplos. Aqui, os valores dos parâmetros escolhidos são os mesmos dos apresentados anteriormente:

- $n = 90$ e $k = 9$ para: AA, EMBR3, IBM, LIGT3, MCD, NKE, POMO4, T, TIF, USIM5
- $n = 100$ e $k = 10$ para: AAPL, ABEV3, BBDC3, CRUZ3, ELET3
- $n = 150$ e $k = 15$ para: ITSA4, KO, MMM, PETR3, SBUX

A tabela 4.5 apresenta um quadro comparativo entre os resultados obtidos pelo *B&H* e pelas abordagens utilizando PAA-PSO-CD e PAA-IDPSO-CD, ambas operando nos dois tipos de estratégia e com retreino baseado na detecção de mudança de conceito. Na primeira coluna, PSO-COMP representa a abordagem PAA-PSO-CD operando apenas na posição comprado, PSO-COMP-VEND, a abordagem PAA-PSO-CD operando tanto em comprado quanto em vendido, IDPSO-COMP representa a abordagem PAA-IDPSO-CD operando apenas na posição comprado, IDPSO-COMP-VEND, a abordagem PAA-IDPSO-CD operando tanto em comprado quanto em vendido. A coluna média e desvio estão relacionadas aos resultados finais do período obtidos nas 50 execuções e a última coluna, tempo, apresenta o tempo gasto, em segundos, para executá-las. Em negrito estão as abordagens que, em média, obtiveram os melhores resultados para cada ação.

Como pode ser observado na tabela 4.5, em relação ao custo computacional, assim como na subseção anterior, utilizar a estratégia comprado-vendido é muito mais custosa que operar apenas em comprado. Aqui, pode-se chegar às mesmas conclusões: a estratégia comprado-vendido é mais custosa por conta das regras de decisão e estrutura da partícula mais complexas e, também, o IDPSO é mais custoso que o PSO.

Tabela 4.5: Comparativo dos valores obtidos pelo PAA-PSO-CD E PAA-IDPSO-CD, para as 50 execuções.

ABORDAGEM	AA			AAPL		
	MÉDIA	DESVIO	TEMPO	MÉDIA	DESVIO	TEMPO
PSO-COMP	489404	206161.9	1142 s	1436009	522372	1136 s
PSO-COMP-VEND	609066.2	283330.7	2447 s	1394993	642038.1	3543 s
IDPSO-COMP	540090.4	242819.1	1599 s	1596357	649101.2	1361 s
IDPSO-COMP-VEND	778641.5	357334.7	2890 s	1422707	777600.7	4567 s
BUY AND HOLD	48673.78			616130.10		
ABORDAGEM	ABEV3			BBDC3		
PSO-COMP	1837473	482601.1	860 s	1470730	345821	987 s
PSO-COMP-VEND	1585719	522500.2	2919 s	1285693	449929.1	2703 s
IDPSO-COMP	1893699	520762.6	898 s	1619027	386093.8	1039 s
IDPSO-COMP-VEND	1647700	827593.2	4167 s	1460514	403524.5	3233 s
BUY AND HOLD	270757.36			264680.18		
ABORDAGEM	CRUZ3			ELET3		
PSO-COMP	1149938	433965.6	1655 s	717330.8	344643.1	1202 s
PSO-COMP-VEND	1156119	474828	4756 s	690051.9	252941.2	2330 s
IDPSO-COMP	1225197	364459.3	1966 s	891082.6	369061	1188 s
IDPSO-COMP-VEND	1162965	459232.1	5692 s	854590.7	371910.6	3231 s
BUY AND HOLD	688834.80			148133.99		
ABORDAGEM	EMBR3			IBM		
PSO-COMP	293670.7	197028.6	1559 s	268197.1	59488.95	1128 s
PSO-COMP-VEND	536459.3	221413.9	2664 s	406375.3	153544	2151 s
IDPSO-COMP	356418.2	246130.6	1983 s	278271.6	70338.35	1339 s
IDPSO-COMP-VEND	639368.6	263303.7	2767 s	496353.2	229404.4	2616 s
BUY AND HOLD	56840.29			187613.85		
ABORDAGEM	ITSA4			KO		
PSO-COMP	567675.4	274207	994 s	213519.3	42633.01	1957 s
PSO-COMP-VEND	586535	269605.2	5219 s	236576.8	50409.59	4610 s
IDPSO-COMP	724969.9	335444.8	1010 s	240527.8	53373.4	2206 s
IDPSO-COMP-VEND	715926.8	398383.8	6145 s	253615.1	60626.66	5245 s
BUY AND HOLD	213471.82			138115.42		
ABORDAGEM	LIGT3			MCD		
PSO-COMP	125816.3	63689.52	2302 s	389552.3	60771.77	1058 s
PSO-COMP-VEND	120440.3	105107.7	6539 s	373309.3	93233.51	2415 s
IDPSO-COMP	154411.2	152346.4	2655 s	417873.9	70084.7	1069 s
IDPSO-COMP-VEND	130734	107345	8731 s	403318.3	106768.4	2859 s
BUY AND HOLD	53718.21			266776.40		
ABORDAGEM	MMM			NKE		
PSO-COMP	188370.3	52302.17	1856 s	402510.8	125390.7	1129 s
PSO-COMP-VEND	301184.9	96831.81	3040 s	441841.4	124563.3	2649 s
IDPSO-COMP	201084.2	66129.43	2082 s	445295.3	118934.5	1183 s
IDPSO-COMP-VEND	324477.1	142021.9	3800 s	496400	164677.9	2859 s
BUY AND HOLD	130155.20			195359.53		

Tabela 4.5: Continuação.

ABORDAGEM	PETR3			POMO4		
PSO-COMP	801069	378612.2	1148 s	1384735	496560.5	966 s
PSO-COMP-VEND	1059541	476111.5	3491 s	1102610	533036.7	4243 s
IDPSO-COMP	1054488	542673.5	1277 s	1598608	510708.7	1029 s
IDPSO-COMP-VEND	1190071	684533.5	4340 s	1143638	683675.9	6445 s
BUY AND HOLD	460065.15			421429.69		
Abordagem	SBUX			T		
PSO-COMP	261084	92960.57	2122 s	261684.4	46405.76	1022 s
PSO-COMP-VEND	312639.1	125532.2	4516 s	260974.8	68266.22	2458 s
IDPSO-COMP	315461.9	105835.4	2225 s	289383	61889.58	1174 s
IDPSO-COMP-VEND	406335	154525.6	5111 s	315074	87365.82	2873 s
BUY AND HOLD	100389.61			140190.06		
Abordagem	TIF			USIM5		
PSO-COMP	579610.5	165257.7	907 s	836235.7	273138.5	964 s
PSO-COMP-VEND	598707.4	320445.6	2296 s	828031.9	412830.9	3103 s
IDPSO-COMP	571076.2	232142.6	1227 s	974663.1	420407.8	1047 s
IDPSO-COMP-VEND	806899.1	355524.4	2687 s	838020.8	499067.4	3683 s
BUY AND HOLD	158109.70			259157.41		

Além disso, a análise dos resultados da tabela 4.5 mostra que o modelo proposto apresentou melhores resultados quando comparado a estratégia *Buy-and-Hold*. Esta consiste em o investidor comprar as ações, mantendo-as em sua carteira de investimentos por um longo período de tempo. O *Buy-and-Hold* é amplamente utilizado como uma referência nos estudos aplicados ao mercado financeiro (TEIXEIRA; OLIVEIRA, 2010). Os testes estatísticos mostraram que todas as abordagens com detecção de mudança de conceito (PSO-COMP, PSO-COMP-VEND, IDPSO-COMP, IDPSO-COMP-VEND) apresentaram resultados estatisticamente melhores que o *B&H* ($p - \text{valor} < 0,05$).

Inicialmente, para facilitar a visualização do que será explicado nos testes estatísticos, são apresentados os *boxplots* dos resultados finais do exercício para todas as ações. As figuras de 4.9 até 4.13 apresentam o *boxplot* dos resultados obtidos pelas variações do PAA-PSO e PAA-IDPSO para as ações. As figuras de 4.9 até 4.13 apresentam os *boxplots* para as ações. O foco, aqui, deve ser na seção “Sem Retreino” das imagens. Nessas imagens, estão os *boxplots* das estratégias sem retreino: PAA-PSO operando apenas em posição comprado e sem retreino (psoCompS), PAA-PSO operando em posição comprado e vendido e sem retreino (psoCompVendS), PAA-IDPSO operando apenas em posição comprado e sem retreino (idpsoCompS), PAA-IDPSO operando em posição comprado e vendido e sem retreino (idpsoCompVendS). Ainda, estão presente os *boxplots* das abordagens com retreino: PAA-PSO operando apenas em posição comprado e com retreino (psoCompC), PAA-PSO operando em posição comprado e vendido e com retreino (psoCompVendC), PAA-IDPSO operando apenas em posição comprado e com retreino (idpsoCompC), PAA-IDPSO operando em posição comprado e vendido e com retreino (idpsoCompVendC). Por fim, estão os *boxplots* da estratégia comprado-vendido com partição de

saldo inicial e com retreino (idpsoCompVendP).

Tabela 4.6: Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO e PAA-IDPSO, na comparação dos modelos com e sem detecção de mudança de conceito, nas 50 execuções, com nível de confiança de 95%.

ABORDAGENS	RESULTADOS DOS TESTES ESTATÍSTICOS
PSO-Comp	A abordagem que utilizou a detecção de mudança de conceito obteve resultados estatisticamente melhores ($p - valor < 0,05$) em todas as ações, quando comparado à mesma abordagem sem detecção
IDPSO-Comp	A abordagem que utilizou a detecção de mudança de conceito obteve resultados estatisticamente melhores ($p - valor < 0,05$) em todas as ações, quando comparado à mesma abordagem sem detecção
PSO-Comp-Vend	A abordagem que utilizou a detecção de mudança de conceito obteve resultados estatisticamente melhores ($p - valor < 0,05$) em todas as ações, quando comparado à mesma abordagem sem detecção
IDPSO-Comp-Vend	A abordagem que utilizou a detecção de mudança de conceito obteve resultados estatisticamente melhores ($p - valor < 0,05$) em todas as ações, quando comparado à mesma abordagem sem detecção

A partir dos *boxplots* (figuras 4.9 a 4.13) e do resultado apresentado na tabela 4.6, as abordagens que utilizaram a detecção de mudança de conceito apresentaram melhores resultado quando comparadas àquelas que não utilizaram a detecção, com 95% de confiança.

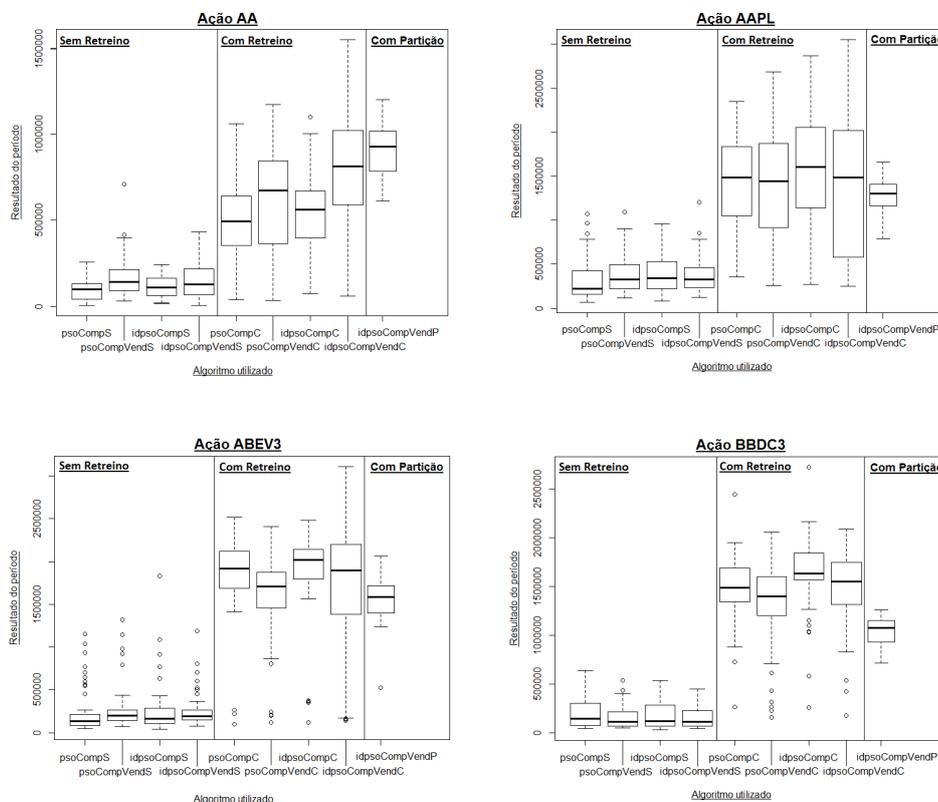


Figura 4.9: *Boxplots* do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações AA, AAPL, ABEV3 e BBDC3.

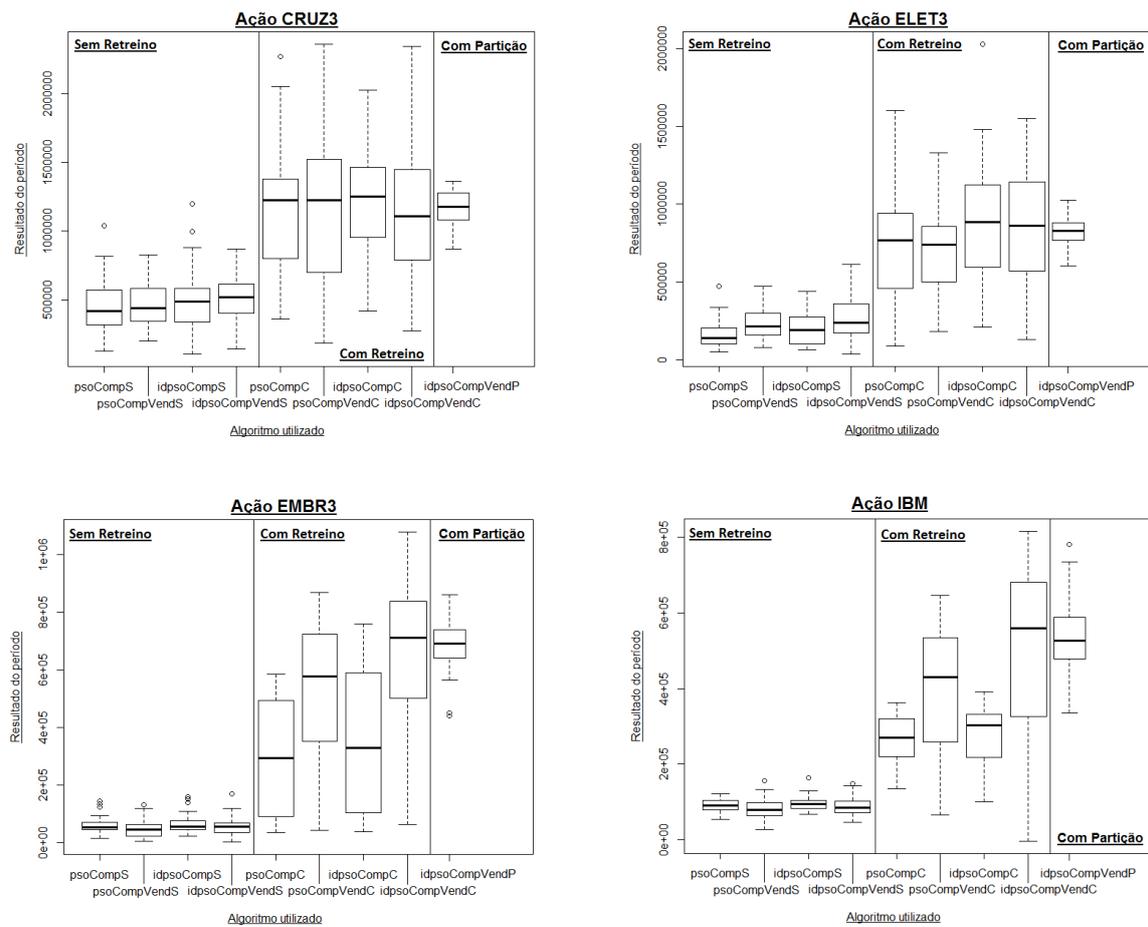


Figura 4.10: *Boxplots* do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações CRUZ3, ELET3, EMBR3 e IBM.

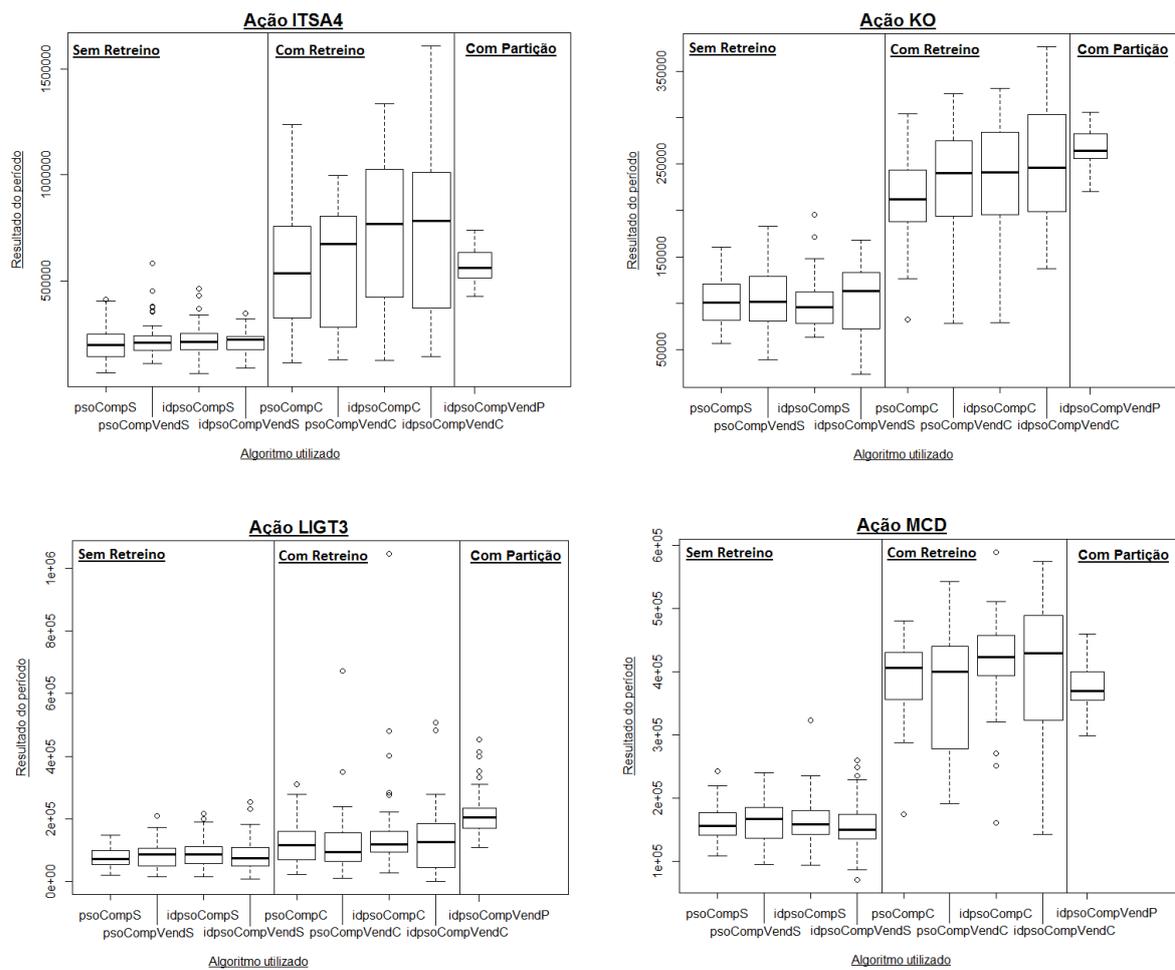


Figura 4.11: *Boxplots* do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações ITSA4, KO, LIGT3 e MCD.

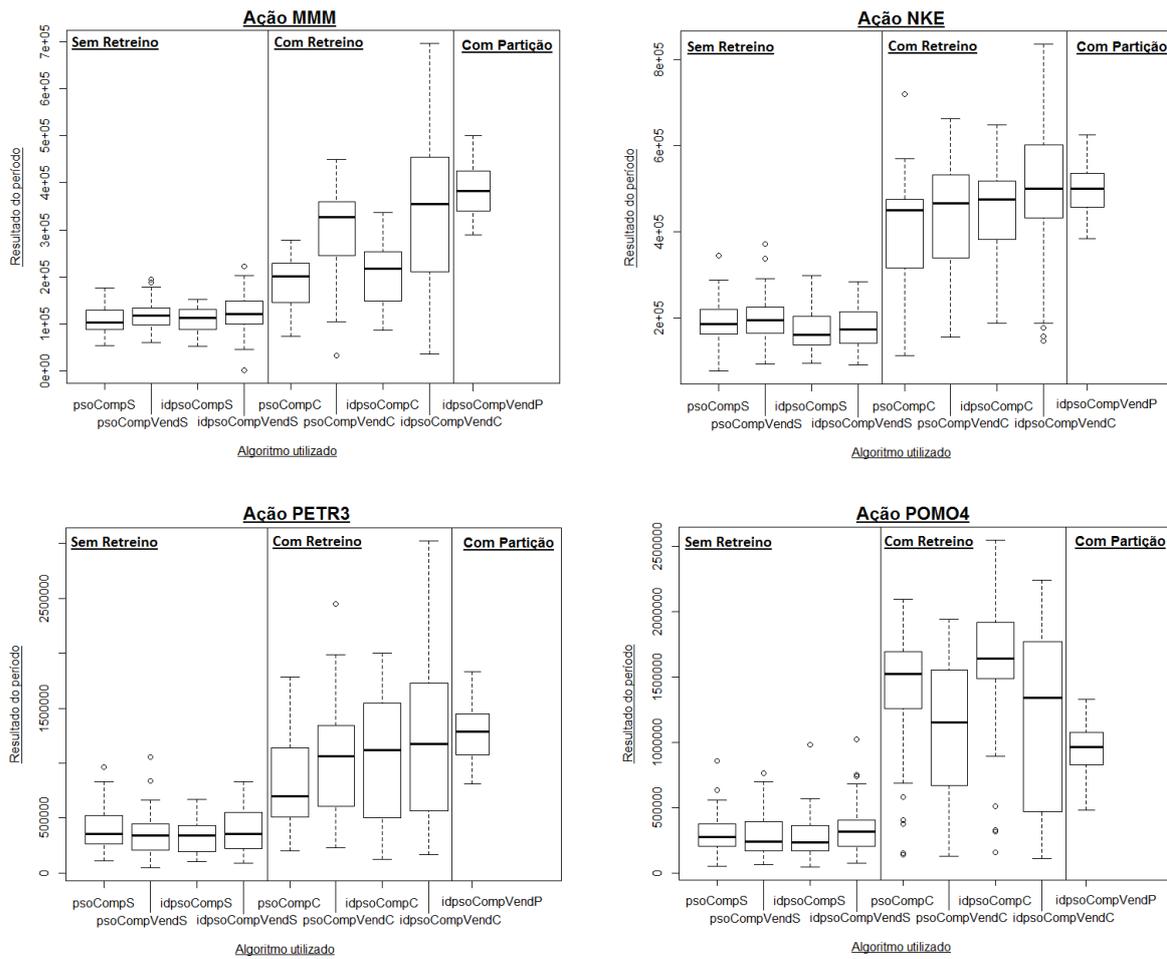


Figura 4.12: *Boxplots* do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações MMM, NKE, PETR3 e POMO4.

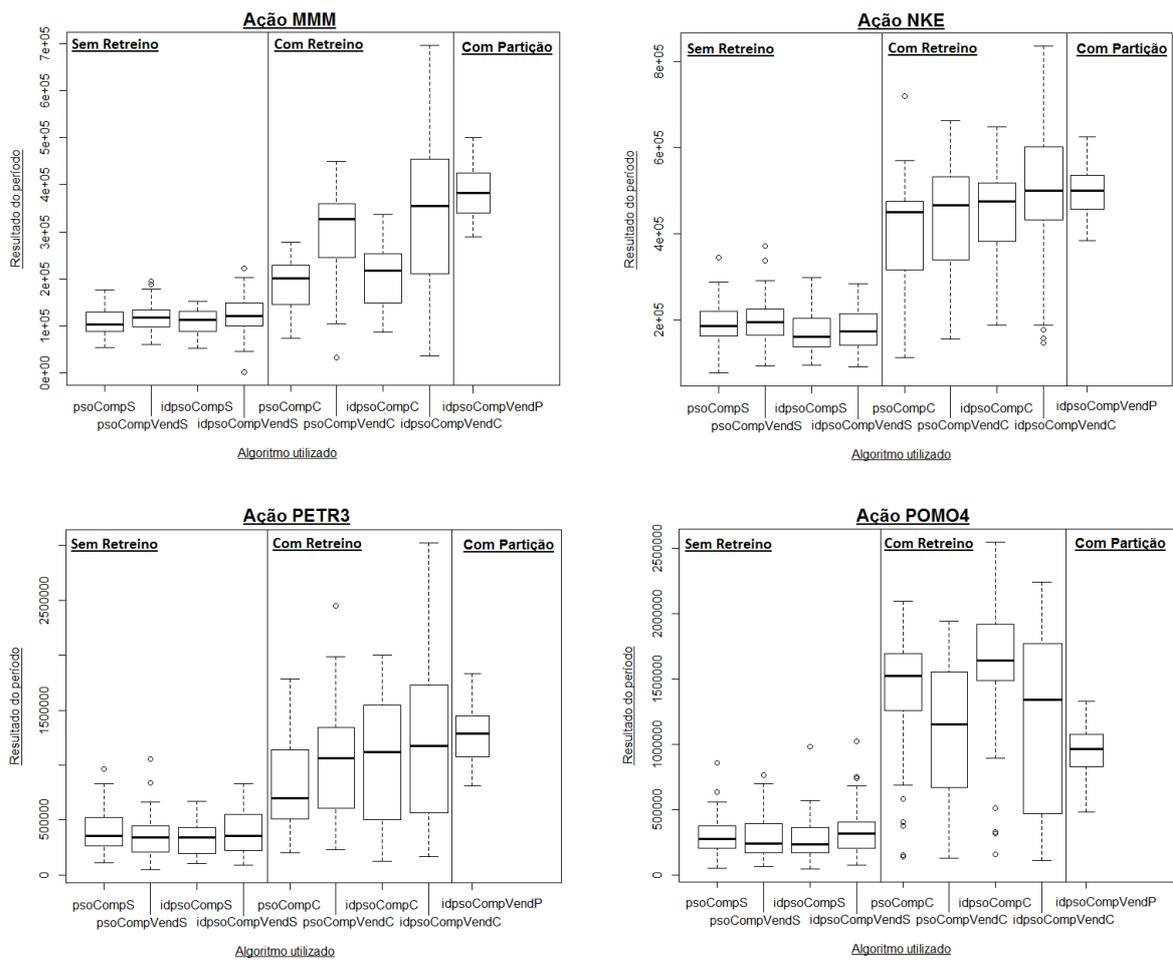


Figura 4.13: *Boxplots* do resultado financeiro final do período obtido pelos modelos PAA-PSO e PAA-IDPSO nas 50 execuções para as ações SBUX, T, TIF e USIM5.

Por sua vez, no comparativo estatístico entre os modelos que utilizaram a detecção de mudança de conceito, a estatística comparativa foi executada em duas etapas: primeiramente, fixando a abordagem utilizada (PAA-PSO ou PAA-IDPSO) e variando as estratégias de operacionalização, na sequência, fixando a estratégias de operacionalização e variando abordagem utilizada. A tabela 4.7 apresenta os resultados para esses testes estatísticos.

Como pode ser observado pela tabela 4.7 na comparação entre PSO-Comp e PSO-Comp-Vend, a abordagem utilizando apenas operações comprado apresentou melhores resultados para as ações ABEV3, BBDC3, POMO4, por sua vez, a abordagem comprado-vendido foi melhor em AA, EMBR3, IBM, KO, MMM, PETR3, SBUX. Em relação à comparação IDPSO-Comp e IDPSO-Comp-Vend, a abordagem utilizando apenas operações comprado apresentou melhores resultados para as ações BBDC3, POMO4, por sua vez, a abordagem comprado-vendido foi melhor em AA, EMBR3, IBM, MMM, SBUX, T, TIF.

Tabela 4.7: Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-PSO-CD e PAA-IDPSO-CD, nas 50 execuções, com nível de confiança de 95%.

ABORDAGENS	RESULTADOS DOS TESTES ESTATÍSTICOS
PSO-Comp X PSO-Comp-Vend	PSO-Comp: resultados estatisticamente melhores ($p - valor < 0,05$) em: ABEV3, BBDC3, POMO4 PSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, EMBR3, IBM, KO, MMM, PETR3, SBUX
IDPSO-Comp X IDPSO-Comp-Vend	IDPSO-Comp resultados estatisticamente melhores ($p - valor < 0,05$) em: BBDC3, POMO4 IDPSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, EMBR3, IBM, MMM, SBUX, T, TIF
PSO-Comp X IDPSO-Comp	IDPSO-Comp: estatisticamente melhor ($p - valor < 0,05$) em: USIM5, ITSA4, BBDC, ELET, KO, MCD, PETR3, POMO4, SBUX, T
PSO-Comp-Vend X IDPSO-Comp-Vend	IDPSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, ITSA4, BBDC3, ELET3, EMBR3, IBM, NKE, SBUX, T, TIF

Um estudo mais aprofundado sobre o comportamento da série original deve ser feito, para a decisão de se usar a estratégia operando apenas em comprado ou pela operante em comprado e vendido. Por exemplo, caso seja observado que as cotações apresentam apenas tendência de subida, deve-se optar pela regra de negócio que opera apenas em posição comprado, pois com esse comportamento o lucro ocorre com a compra a valor menor e posterior venda a um valor maior. Num outro cenário, em que haja alternâncias entre tendências de subida e de descida, deve-se optar pelo uso da regra de negócio que opera em comprado e vendido, pois como ela pode gerar lucro tanto na tendência de subida (operando comprado) quanto na tendência de descida (operando vendido), tem maior capacidade de se adequar a esse tipo de cenário. Como consequência dessa maior capacidade de adequação às variações de tendência, a estratégia comprado-vendido apresentou melhores resultado em mais ações, para ambas as abordagens utilizando PAA-PSO-CD e PAAA-IDPSO-CD.

Na comparação em utilizar a otimização do PSO ou do IDPSO, o IDPSO apresentou uma melhor otimização quando comparado ao PSO. O que era de se esperar por conta de sua melhor dinamicidade e maior capacidade de exploração e exploração. Isso pode ser verificado através da melhor mineração que o PAA-IDPSO-CD obteve quando comparado ao PAA-PSO-CD. Como pode ser visto, tanto nos *boxplots* quanto na tabela 4.7, o IDPSO apresentou resultados melhores ou iguais ao PSO, para todas as ações. Na estratégia operando apenas em posição comprado, o IDPSO apresentou resultados estatisticamente melhores nas ações USIM5, ITSA4, BBDC, ELET, KO, MCD, PETR3, POMO4, SBUX, T. Por sua vez, na estratégia operando em comprado-vendido, as ações em que o PAA-IDPSO-CD apresentou melhores resultados foram AA, ITSA4, BBDC3, ELET3, EMBR3, IBM, NKE, SBUX, T, TIF.

Com isso, pode-se afirmar que, para a abordagem de mineração de dados proposta, a utilização da técnica de detecção de mudança de conceito com retreino proposta apresentou-se como um diferencial. Por conta dela, os modelos que utilizaram-na apresentaram resultado significativamente melhores quando comparadas às suas respectivas abordagens sem a detecção. Quanto ao algoritmo de otimização, a opção pelo IDPSO apresenta um resultado final do período melhor quando comparado ao PSO, no entanto, a um custo computacional maior. Com isso, o usuário deve definir sua prioridade: um modelo com melhores resultados ou um modelo de menor custo computacional. Por fim, a estratégia comprado-vendido apresentou melhores resultado em um maior número de ações. Como pode ser observado na tabela 4.7, o IDPSO-Comprado-Vendido apresentou resultados estatisticamente melhores em sete ações, o IDPSO-Comprado apresentou em duas ações e em onze ações não houve diferença significativa. No entanto, um estudo mais aprofundado do comportamento da série deve ser feito para se ter mais informações sobre qual das estratégias deve ser utilizada.

4.2.4 Comparação PAA-IDPSO-CD comprado-vendido com e sem partição de saldo inicial

Como apresentado na subseção anterior, a técnica PAA-IDPSO-CD operando nas estratégias comprado-vendido e usando retreino baseado na detecção de mudança de conceito foi a que apresentou resultados estatisticamente melhores. No entanto, como pode ser visto nos *boxplots* de 4.9 até 4.13 e na tabela 4.5, esse modelo apresenta uma variância muito grande em seus resultados. Visando reduzir esta variância, propusemos, na seção 4.4.2, um método baseado na partição do saldo inicial. Nessa subseção são feitas as comparações entre PAA-IDPSO-CD comprado-vendido com e sem partição de saldo inicial, para isso utilizou-se dez partições do saldo inicial $n = 10$. Como o saldo inicial, $SI = \$100.000,00$ e $n = 10$, então $SI_x = 100.000,00/10 = \$10.000,00$.

A escolha por $n = 10$ decorre dos seguintes fatos: 1) com n muito pequeno, fica mais difícil de ocorrer a compensação entre a perda de um modelo e ganho de outro e isso manterá a variância do modelo grande; 2) caso n seja muito grande, o saldo inicial de cada modelo SI_x será

muito pequeno, e conseqüentemente difícil de operar no mercado de ações. O $SI_x = \$10.000,00$, apresentou-se como um bom valor para realizar as operações.

As variáveis a serem comparadas são, novamente, o resultado financeiro final do período e o tempo de execução. São utilizados os mesmos valores de tamanho de janela (n) e o número de segmentos (k) e todas as ações definidas anteriormente com o objetivo de obter resultados mais amplos.

Os resultados apresentados nas figuras 4.9 a 4.13 e, também, na tabela 4.8 - que traz um comparativo entre os valores obtidos pelo PAA-IDPSO em suas configurações comprado, comprado-vendido e comprado-vendido com partição de saldo inicial, todas com retreino a partir da detecção de mudança de conceito.

Na primeira coluna da tabela 4.8, *B&H* representa a abordagem *Buy and Hold*, IDPSO-COMP-VEND representa a abordagem PAA-IDPSO-CD operando em comprado e vendido sem partição do saldo inicial e IDPSO-PARTIÇÃO representa a abordagem PAA-IDPSO-CD operando em comprado e vendido com partição do saldo inicial. A coluna média e desvio estão relacionadas aos resultados finais do período obtidos nas 50 execuções e a última coluna, tempo, apresenta o tempo gasto, em segundos, para executá-las. Em negrito estão as abordagens que obtiveram os melhores resultados médios, para cada ação.

Em relação ao custo computacional, como pode ser observado, utilizar o particionamento do saldo inicial é muito mais custoso que operar sem ele. Esse fato decorre de nessa nova metodologia, cada iteração representa a execução de 10 modelos diferentes. Então, caso o investidor precise de uma resposta mais rápida, deve utilizar a metodologia sem particionamento.

A partir da tabela 4.8 pode ser visualizado que as duas abordagens apresentaram melhores resultados quando comparado a estratégia *Buy-and-Hold*. Ao executar os testes estatísticos, as duas abordagens (IDPSO-COMP-VEND, IDPSO-PARTIÇÃO) apresentaram resultados estatisticamente melhores que o *B&H* ($p - valor < 0,05$).

Também, em decorrência da tabela 4.8, pode ser observado que a metodologia proposta atingiu seus objetivos. Com isso, foi reduzida consideravelmente a variância para todas as ações e, conseqüentemente, o risco das operações.

A tabela 4.9 apresenta os resultados para os testes estatísticos entre os modelos com e sem partição de saldo inicial, para a estratégia operando em comprado-vendido. Como pode ser observado, o IDPSO-Comp-Vend, sem partição, apresentou resultados estatisticamente melhores para as ações BBDC3, ITSA4. Por sua vez, IDPSO-Partição, com partição do saldo inicial, obteve resultados significativamente melhores para as ações AA, LIGT3, MMM. Na maior parte dos casos, ou seja, nas outras 15 ações os resultados não apresentaram diferença significativa. Mesmo assim, a metodologia de investimento apresentada conseguiu atingir seu objetivo, uma vez que conseguiu reduzir consideravelmente a variância do modelo.

Assim, a decisão de usar a metodologia de investimento com ou sem partição do saldo inicial irá depender das características do próprio investidor. Caso o investidor esteja disposto a correr maiores riscos objetivando maiores lucros potenciais então deve utilizar a metodologia

Tabela 4.8: Comparativo dos valores obtidos pelo PAA-IDPSO-CD comprado-vendido com e sem partição de saldo inicial, para as 50 execuções.

ABORDAGEM	AA			AAPL		
	MÉDIA	DESVIO	TEMPO	MÉDIA	DESVIO	TEMPO
IDPSO-COMP-VEND	778641.5	357334.7	2890 s	1422707	777600.7	4567 s
IDPSO-PARTIÇÃO	911341.2	149784.3	29535 s	1281763	206419.4	41042 s
BUY AND HOLD	48673.78			616130.10		
ABORDAGEM	ABEV3			BBDC3		
IDPSO-COMP-VEND	1647700	827593.2	4167 s	1460514	403524.5	3233 s
IDPSO-PARTIÇÃO	1568972	247313.3	33561 s	1038256	137404.7	32238 s
BUY AND HOLD	270757.36			264680.18		
ABORDAGEM	CRUZ3			ELET3		
IDPSO-COMP-VEND	1162965	459232.1	5692 s	854590.7	371910.6	3231 s
IDPSO-PARTIÇÃO	1168453	127500.3	49484 s	824961.3	93162.6	28276 s
BUY AND HOLD	688834.80			148133.99		
ABORDAGEM	EMBR3			IBM		
IDPSO-COMP-VEND	639368.6	263303.7	2767 s	496353.2	229404.4	2616 s
IDPSO-PARTIÇÃO	688491.5	89433.33	27273 s	532284.3	88457.07	23942 s
BUY AND HOLD	56840.29			187613.85		
ABORDAGEM	ITSA4			KO		
IDPSO-COMP-VEND	715926.8	398383.8	6145 s	253615.1	60626.66	5245 s
IDPSO-PARTIÇÃO	572877.5	80063.1	51959 s	266846.9	17937.69	46899 s
BUY AND HOLD	213471.82			138115.42		
ABORDAGEM	LIGT3			MCD		
IDPSO-COMP-VEND	130734	107345	8731 s	403318.3	106768.4	2859 s
IDPSO-PARTIÇÃO	218545.6	73326.3	46586 s	374336.6	33389.2	28531 s
BUY AND HOLD	53718.21			266776.40		
ABORDAGEM	MMM			NKE		
IDPSO-COMP-VEND	324477.1	142021.9	3800 s	496400	164677.9	2859 s
IDPSO-PARTIÇÃO	385586.5	54851.99	38445 s	499138.1	56126.8	27757 s
BUY AND HOLD	130155.20			195359.53		
ABORDAGEM	PETR3			POMO4		
IDPSO-COMP-VEND	1190071	684533.5	4340 s	1143638	683675.9	6445 s
IDPSO-PARTIÇÃO	1278281	239735	40165 s	968622.1	183969.7	44249 s
BUY AND HOLD	460065.15			421429.69		
Abordagem	SBUX			T		
IDPSO-COMP-VEND	406335	154525.6	5111 s	315074	87365.82	2873 s
IDPSO-PARTIÇÃO	386452.7	43397.27	39847 s	306427.4	35997.6	26165 s
BUY AND HOLD	100389.61			140190.06		
Abordagem	TIF			USIM5		
IDPSO-COMP-VEND	806899.1	355524.4	2687 s	838020.8	499067.4	3683 s
IDPSO-PARTIÇÃO	801987.1	140185.7	25542 s	945278.4	220251.9	33297 s
BUY AND HOLD	158109.70			259157.41		

Tabela 4.9: Testes estatísticos para o comparativo dos resultados obtidos pelos modelos PAA-IDPSO operando em comprado-vendido com e sem partição de saldo inicial, nas 50 execuções, com nível de confiança de 95%.

ABORDAGENS	RESULTADOS DOS TESTES ESTATÍSTICOS
IDPSO-Comp-Vend X IDPSO-Partição	IDPSO-Comp-Vend: resultados estatisticamente melhores ($p - valor < 0,05$) em: BBDC3, ITSA4 IDPSO-Partição: resultados estatisticamente melhores ($p - valor < 0,05$) em: AA, LIGT3, MMM

sem partição, que pode resultar em resultados máximos maiores. Por outro lado, caso o investidor seja mais cauteloso, ele deve optar pela partição do saldo inicial, por apresentar uma variância menor.

Na Figura 4.14 a coluna da esquerda apresenta os resultados obtidos pelo PAA-IDPSO-CD operando em posições comprado-vendido e sem partição do saldo inicial, a coluna da direita apresenta o mesmo modelo com o uso da partição do saldo inicial, ambos para a ação AA. Como pode ser observado, devido ao fato de o IDPSO ser um algoritmo de otimização estocástico, ao rodar o PAA-IDPSO-CD, o investidor pode obter o valor máximo, mas também pode obter o valor mínimo. No caso do uso da abordagem com partição, o modelo obtém resultados mais próximo, ou seja, o risco é menor.

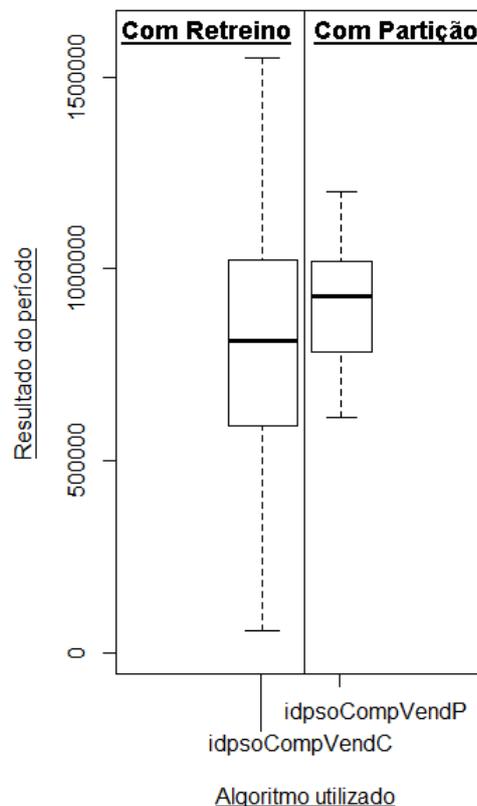


Figura 4.14: Boxplot do PAA-IDPSO-CD operando em posições comprado-vendido, com e sem partição do saldo inicial, nas 50 execuções para a ação AA.

4.3 Considerações finais

Nesse capítulo, inicialmente foram apresentados os resultados comparativos entre o PAA-PSO e o SAX-GA para mostrar a viabilidade do modelo inicial. Posteriormente foram mostrados os resultados das diversas versões até chegar ao modelo final, o PAA-IDPSO-CD. Essas abordagens foram comparadas entre si e com a estratégia *Buy and Hold*. Para ratificar os resultados obtidos, ainda foram realizados testes estatísticos, Testes *t* Pareados para demonstrar a diferença significativa entre as diversas abordagens.

A partir dos resultados, pode ser observado que a utilização da detecção de mudança de conceito apresentou uma melhoria significativa nos resultados obtidos quando comparados a não utilização desta estratégia proposta. Além disso, o IDPSO apresentou resultados melhores que o PSO, para o modelo proposto. Com relação à regra de decisão a ser utilizada, a estratégia comprado-vendido apresentou melhores resultados em um maior número de ações. No entanto, estudos mais aprofundados devem ser realizados para se ratificar qual das estratégias deve ser utilizada em cada ação.

Por fim, a depender do grau de risco que o investidor esteja disposto a aceitar ele deve optar por uma das metodologia de investimento apresentadas. Caso ele aceite um maior risco, deve utilizar a metodologia sem partição do saldo inicial, pois o lucro máximo potencial é maior. Por outro lado, caso ele seja mais cauteloso, ele deve optar pela partição do saldo inicial. Em relação ao custo computacional, caso o investidor precise de uma resposta mais rápida, deve utilizar a metodologia sem particionamento.

5

Conclusão

O presente estudo apresentou uma nova abordagem, o PAA-IDPSO-CD, para descoberta de padrões em séries temporais financeiras e otimização das operações de compras e vendas a serem feitas no mercado de ações. Para isso, são combinadas as técnicas Aproximação por Valor Agregado de Segmento (PAA), como método de representação de séries temporais, com o algoritmo de Otimização por Enxame de Partículas Auto Adaptativa (IDPSO). Por fim, é proposta uma métrica de detecção de mudança de conceito, para a identificação dos momentos que o sistema deve ser retreinado durante o uso do sistema (fase de teste). Os padrões encontrados são, então, utilizados junto à regras de decisão propostas com o objetivo de servir como um sistema de suporte de decisão responsável por automatizar as operações feitas no mercado de ações.

A partir dos resultados apresentados nos experimentos, pode ser observado que o modelo proposto demonstrou-se viável e, quando comparado ao SAX-GA, apresenta vantagens como: ser menos custoso computacionalmente, necessidade de menos parâmetros e obter lucros iguais ou superiores. Além disso, a utilização da detecção de mudança de conceito apresentou uma melhoria significativa nos resultados obtidos em comparação aos modelos que não utilizam tal detecção para todas as ações. Os modelos operando com detecção de mudança de conceito também obtiveram melhores resultados quando comparados a estratégia Buy and Hold.

Em relação ao algoritmo de otimização utilizado, o IDPSO apresentou melhores resultados quando comparado ao PSO. Nessa comparação, para a estratégia operando apenas em posição comprado, o PAA-IDPSO-CD apresentou resultados estatisticamente melhores que o PAA-PSO-CD em dez ações e em dez não houve diferença significativa. Por sua vez na estratégia operando em posições comprado e vendido, o PAA-IDPSO-CD também apresentou resultados estatisticamente melhores em dez ações e em dez não houve diferença significativa, quando comparado ao PAA-PSO-CD.

Com relação à estratégia de investimento a ser utilizada, a estratégia comprado-vendido apresentou melhores resultados quando comparados à estratégia operando apenas em posição comprado. Nessa comparação, O PAA-IDPSO-CD operando comprado-vendido apresentou resultados estatisticamente melhores em sete ações, o PAA-IDPSO-CD operando apenas em posições comprado foi estatisticamente superior em duas ações e em onze ações não houve

diferença significativa. Assim, a estratégia operando em posições comprado-vendido apresentou uma maior capacidade de adequação aos diversos comportamentos de séries temporais financeiras do mercado de ações.

Por fim, a depender do grau de risco que o investidor esteja disposto a aceitar ele pode optar por uma das metodologia de investimento apresentadas. Caso ele aceite um maior risco, deve utilizar a metodologia sem partição do saldo inicial, pois o lucro máximo potencial é maior. Por outro lado, caso ele seja mais cauteloso e aceite apenas um risco menor, ele deve optar pela partição do saldo inicial.

5.1 Contribuições

Entre as contribuições do presente estudo estão as seguintes:

- Proposição de uma nova métrica para a detecção de mudança de conceito para séries temporais financeiras baseada no *Máximo Drawdown*.
- Desenvolvimento de um modelo de mineração de séries temporais baseado em Aproximação por Valor Agregado de Segmento (PAA), no algoritmo de otimização Otimização por Enxame de Partículas Auto Adaptativa (IDPSO) e na detecção de mudança de conceitos na série temporais com base na nova métrica proposta.
- Desenvolvimento de um sistema de suporte de decisão responsável pela automação das operações em mercado de ações.
- Publicação de um artigo com parte dos resultados obtidos na IJCNN 2015 (International Joint Conference on Neural Networks) (SOUZA; BRASILEIRO; OLIVEIRA, 2015)

5.2 Restrições da abordagem proposta

Entre as limitações do presente estudo estão as seguintes:

- Os parâmetros tamanho da janela e quantidade de segmentos, utilizados pelo Aproximação por Valor Agregado de Segmento (PAA) são escolhidos pelo usuário. Isto é, eles não são valores otimizados.
- Em decorrência da redução da dimensionalidade decorrente do PAA, pode haver perda de informações importantes para a análise das séries temporais, como, por exemplo, valores extremos e informações sobre a dispersão das séries temporais.

5.3 Trabalhos futuros

Como trabalhos futuros podem ser citados os seguintes passos:

- Otimizar os parâmetros tamanho da janela e quantidade de segmentos, utilizados pelo Aproximação por Valor Agregado de Segmento (PAA).
- Utilizar outros algoritmos de otimização, como, por exemplo, algoritmos multiobjetivos.
- Utilizar novas regras de negócio mais sofisticadas.
- Desenvolver método que trabalhe com portfólios.

Referências

ABDUAL-SALAM, M. E.; ABDUL-KADER, H. M.; ABDEL-WAHED, W. F. Comparative study between Differential Evolution and Particle Swarm Optimization algorithms in training of feed-forward neural network for stock price prediction. In: INTERNATIONAL CONFERENCE ON INFORMATICS AND SYSTEMS (INFOS), 7. **Anais...** [S.l.: s.n.], 2010. p.1–8.

ARGIDDI, R. V.; APTE, S. S. Analysis of Fragment Mining on Indian Financial Market. **International Journal of Computer Science and Information Technologies (IJCSIT)**, [S.l.], v.5, n.5, p.6670–6675, 2014.

ASADI, S. et al. Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction. **Knowledge-Based Systems**, [S.l.], v.35, n.0, p.245 – 258, 2012.

BACH, S.; MALOOF, M. Paired Learners for Concept Drift. In: DATA MINING, 2008. ICDM '08. EIGHTH IEEE INTERNATIONAL CONFERENCE ON. **Anais...** [S.l.: s.n.], 2008. p.23–32.

BAENA-GARCIA, M. et al. Early drift detection method. In: IN FOURTH INTERNATIONAL WORKSHOP ON KNOWLEDGE DISCOVERY FROM DATA STREAMS (IWKDDS'06). **Anais...** [S.l.: s.n.], 2006. p.77–86.

BAGHERI, A.; PEYHANI, H. M.; AKBARI, M. Financial forecasting using {ANFIS} networks with Quantum-behaved Particle Swarm Optimization. **Expert Systems with Applications**, [S.l.], v.41, n.14, p.6235 – 6250, 2014.

BANERJEE, M.; PAL, N. R. Feature selection with SVD entropy: some modification and extension. **Information Sciences**, [S.l.], v.264, n.0, p.118 – 134, 2014.

BIFET, A.; GAVALDÀ, R. Learning from Time-Changing Data with Adaptive Windowing. In: SEVENTH SIAM INTERNATIONAL CONFERENCE ON DATA MINING, APRIL 26-28, 2007, MINNEAPOLIS, MINNESOTA, USA. **Proceedings...** [S.l.: s.n.], 2007. p.443–448.

BOX, G. E.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2011. v.734.

BRASILEIRO, R. C. et al. Automatic method for stock trading combining technical analysis and the Artificial Bee Colony Algorithm. In: IEEE CONGRESS ON EVOLUTIONARY COMPUTATION (CEC). **Anais...** [S.l.: s.n.], 2013. p.1810–1817.

BULKOWSKI, T. N. **Encyclopedia of Chart Patterns**. [S.l.]: John Wiley and Sons, Inc, 2005. n.2.

CANELAS, A.; NEVES, R.; HORTA, N. A New SAX-GA Methodology Applied to Investment Strategies Optimization. In: ANNUAL CONFERENCE ON GENETIC AND EVOLUTIONARY COMPUTATION, 14., New York, NY, USA. **Proceedings...** ACM, 2012. p.1055–1062. (GECCO '12).

- CANELAS, A.; NEVES, R.; HORTA, N. A SAX-GA approach to evolve investment strategies on financial markets based on pattern discovery techniques. **Expert Systems with Applications**, [S.l.], v.40, n.5, p.1579 – 1590, 2013.
- CARAIANI, P. The predictive power of singular value decomposition entropy for stock market dynamics. **Physica A: Statistical Mechanics and its Applications**, [S.l.], v.393, n.0, p.571 – 578, 2014.
- CHAKRABARTI, K. et al. Locally Adaptive Dimensionality Reduction for Indexing Large Time Series Databases. **ACM Trans. Database Syst.**, [S.l.], v.27, n.2, p.188–228, June 2002.
- CHEN, M.-Y.; CHEN, B.-T. Online fuzzy time series analysis based on entropy discretization and a Fast Fourier Transform. **Applied Soft Computing**, [S.l.], v.14, Part B, n.0, p.156 – 166, 2014.
- CHEN, W. Artificial bee colony algorithm for constrained possibilistic portfolio optimization problem. **Physica A: Statistical Mechanics and its Applications**, [S.l.], v.429, n.0, p.125 – 139, 2015.
- COWPERTWAIT, P. S.; METCALFE, A. V. **Introductory time series with R**. [S.l.]: Springer Science & Business Media, 2009.
- EBERHART, R. C.; SHI, Y. Comparison between genetic algorithms and particle swarm optimization. In: PORTO, V. et al. (Ed.). **Evolutionary Programming VII**. [S.l.]: Springer Berlin Heidelberg, 1998. p.611–616. (Lecture Notes in Computer Science, v.1447).
- EVANS, C.; PAPPAS, K.; XHAFA, F. Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. **Mathematical and Computer Modelling**, [S.l.], v.58, n.5, p.1249–1266, 2013.
- FISCHER, R.; FISCHER, J. **Candlesticks, Fibonacci, and Chart Pattern Trading Tools: a synergistic strategy to enhance profits and reduce risk**. [S.l.]: John Wiley and Sons, Inc, 2003.
- FU, T. chung. A review on time series data mining. **Engineering Applications of Artificial Intelligence**, [S.l.], v.24, n.1, p.164 – 181, 2011.
- FUAD, M. M. M.; MARTEAU, P.-F. Towards a faster symbolic aggregate approximation method. **arXiv preprint arXiv:1301.5871**, [S.l.], 2013.
- GAMA, J. A survey on learning from data streams: current and future trends. **Progress in Artificial Intelligence**, [S.l.], v.1, n.1, p.45–55, 2012.
- GAMA, J. et al. Learning with Drift Detection. In: **Advances in Artificial Intelligence – SBIA 2004**. [S.l.]: Springer Berlin Heidelberg, 2004. v.3171, p.286–295.
- GAMA, J. et al. A Survey on Concept Drift Adaptation. **ACM Comput. Surv.**, [S.l.], v.46, n.4, p.44:1–44:37, Mar. 2014.
- GORGULHO, A.; NEVES, R.; HORTA, N. Applying a GA kernel on optimizing technical analysis rules for stock picking and portfolio composition. **Expert Systems with Applications**, [S.l.], v.38, n.11, p.14072 – 14085, 2011.

- GUO-QIANG, X. The optimization of share price prediction model based on support vector machine. In: INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION AND SYSTEMS ENGINEERING (CASE). **Anais...** [S.l.: s.n.], 2011. p.1–4.
- HASSAN, R. et al. A Comparison of Particle Swarm Optimization and the Genetic Algorithm. In: AIAA/ASME/ASCE/AHS/ASC STRUCTURES, STRUCTURAL DYNAMICS AND MATERIALS CONFERENCE, 46. **Proceedings...** American Institute of Aeronautics and Astronautics, 2005.
- HUANG, C.-F. A hybrid stock selection model using genetic algorithms and support vector regression. **Applied Soft Computing**, [S.l.], v.12, n.2, p.807–818, 2012.
- HUANG, J.; ZHU, W.; RUAN, X. Option pricing using the fast Fourier transform under the double exponential jump model with stochastic volatility and stochastic intensity. **Journal of Computational and Applied Mathematics**, [S.l.], v.263, n.0, p.152 – 159, 2014.
- JOHANSEN, A. Characterization of large price variations in financial markets. **Physica A: Statistical Mechanics and its Applications**, [S.l.], v.324, n.1–2, p.157 – 166, 2003.
- KANNAN, K. S. et al. Financial Stock Market Forecast using Data Mining Techniques. In: INTERNATIONAL MULTICONFERENCE OF ENGINEERS AND COMPUTER SCIENTIST 2010. **Anais...** [S.l.: s.n.], 2010. (IMECS 2010, v.1).
- KAZEM, A. et al. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. **Applied Soft Computing**, [S.l.], v.13, n.2, p.947 – 958, 2013.
- KENNEDY, J.; EBERHART, R. C. Particle swarm optimization. In: IEEE INTERNATIONAL CONFERENCE ON NEURAL NETWORKS. **Proceedings...** [S.l.: s.n.], 1995. p.1942–1948.
- KEOGH, E. et al. Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. **Knowledge and Information Systems**, [S.l.], v.3, n.3, p.263–286, 2001.
- LI, X.; DENG, Z.; LUO, J. Trading strategy design in financial investment through a turning points prediction scheme. **Expert Systems with Applications**, [S.l.], v.36, n.4, p.7818–7826, 2009.
- LI, Z. et al. Comparative Research on Particle Swarm Optimization and Genetic Algorithm. **Computer and Information Science**, [S.l.], v.3, n.1, p.120–127, 2010.
- LIN, J. et al. A Symbolic Representation of Time Series, with Implications for Streaming Algorithms. In: ACM SIGMOD WORKSHOP ON RESEARCH ISSUES IN DATA MINING AND KNOWLEDGE DISCOVERY, 8., New York, NY, USA. **Proceedings...** ACM, 2003. p.2–11. (DMKD '03).
- LIN, J. et al. Experiencing SAX: a novel symbolic representation of time series. **Data Mining and Knowledge Discovery**, [S.l.], v.15, n.2, p.107–144, 2007.
- LKHAGVA, B.; SUZUKI, Y.; KAWAGOE, K. Extended SAX: extension of symbolic aggregate approximation for financial time series data representation. **The Institute of Electronics, Information and Communication Engineers - (IEICE)**, [S.l.], 2006.
- MAJHI, R. et al. On the development of improved adaptive models for efficient prediction of stock indices using clonal-PSO (CPSO) and PSO techniques. **International Journal of Business Forecasting and Marketing Intelligence**, [S.l.], v.1, n.1, p.50 – 67, 2008.

- MAJHI, R. et al. Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and {BFO} based techniques. **Expert Systems with Applications**, [S.l.], v.36, n.6, p.10097 – 10104, 2009.
- MONTGOMERY, D.; RUNGER, G. **Estatística Aplicada e Probabilidade Para Engenheiros**. [S.l.]: LTC, 2012. v.5.
- MURPHY, J. J. **Technical analysis of the financial markets**. Fishkill, N.Y.: New York Institute of Finance, 1999.
- NISHIDA, K.; YAMAUCHI, K. Detecting Concept Drift Using Statistical Testing. In: **Discovery Science**. [S.l.]: Springer Berlin Heidelberg, 2007. v.4755, p.264–269.
- PULIDO, M.; MELIN, P.; CASTILLO, O. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange. **Information Sciences**, [S.l.], v.280, p.188–204, 2014.
- ROSS, G. J. et al. Exponentially weighted moving average charts for detecting concept drift. **Pattern Recognition Letters**, [S.l.], v.33, n.2, p.191 – 198, 2012.
- SOUZA, V. L.; BRASILEIRO, R. C.; OLIVEIRA, A. L. A PAA-PSO technique for investment strategies in the financial market. In: NEURAL NETWORKS (IJCNN), 2015 INTERNATIONAL JOINT CONFERENCE ON. **Anais...** [S.l.: s.n.], 2015. p.1–8.
- SUN, Y. et al. An improvement of symbolic aggregate approximation distance measure for time series. **Neurocomputing**, [S.l.], v.138, n.0, p.189 – 198, 2014.
- TEIXEIRA, L. A.; OLIVEIRA, A. L. A method for automatic stock trading combining technical analysis and nearest neighbor classification. **Expert Systems with Applications**, [S.l.], v.37, n.10, p.6885 – 6890, 2010.
- THONGMEE, T. et al. Finding Strong Relationships of stock prices using blockwise symbolic representation with dynamic time warping. In: INNOVATIONS IN INTELLIGENT SYSTEMS AND APPLICATIONS (INISTA) PROCEEDINGS, 2014 IEEE INTERNATIONAL SYMPOSIUM ON. **Anais...** [S.l.: s.n.], 2014. p.104–109.
- TSINASLANIDIS, P. E.; KUGIUMTZIS, D. A prediction scheme using perceptually important points and dynamic time warping. **Expert Systems with Applications**, [S.l.], v.41, n.15, p.6848 – 6860, 2014.
- TUNG, W. L.; QUEK, C. Financial volatility trading using a self-organising neural-fuzzy semantic network and option straddle-based approach. **Expert Systems with Applications**, [S.l.], v.38, n.5, p.4668–4688, 2011.
- VANSTONE, B.; FINNIE, G. An Empirical Methodology for Developing Stockmarket Trading Systems Using Artificial Neural Networks. **Expert Systems with Applications**, [S.l.], p.6668–6680, 2009.
- XIONG, T. et al. Forecasting interval time series using a fully complex-valued {RBF} neural network with {DPSO} and {PSO} algorithms. **Information Sciences**, [S.l.], v.305, n.0, p.77 – 92, 2015.

ZHANG, Y.; XIONG, X.; ZHANG, Q. An Improved Self-Adaptive PSO Algorithm with Detection Function for Multimodal Function Optimization Problems. **Mathematical Problems in Engineering**, [S.l.], v.2013, 2013.

ZHU, H. et al. Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. **Expert Systems with Applications**, [S.l.], v.38, n.8, p.10161 – 10169, 2011.

Apêndice



Séries temporais das ações utilizadas

As figuras nesse apêndice apresentam as séries temporais das ações utilizadas pelo presente estudo.

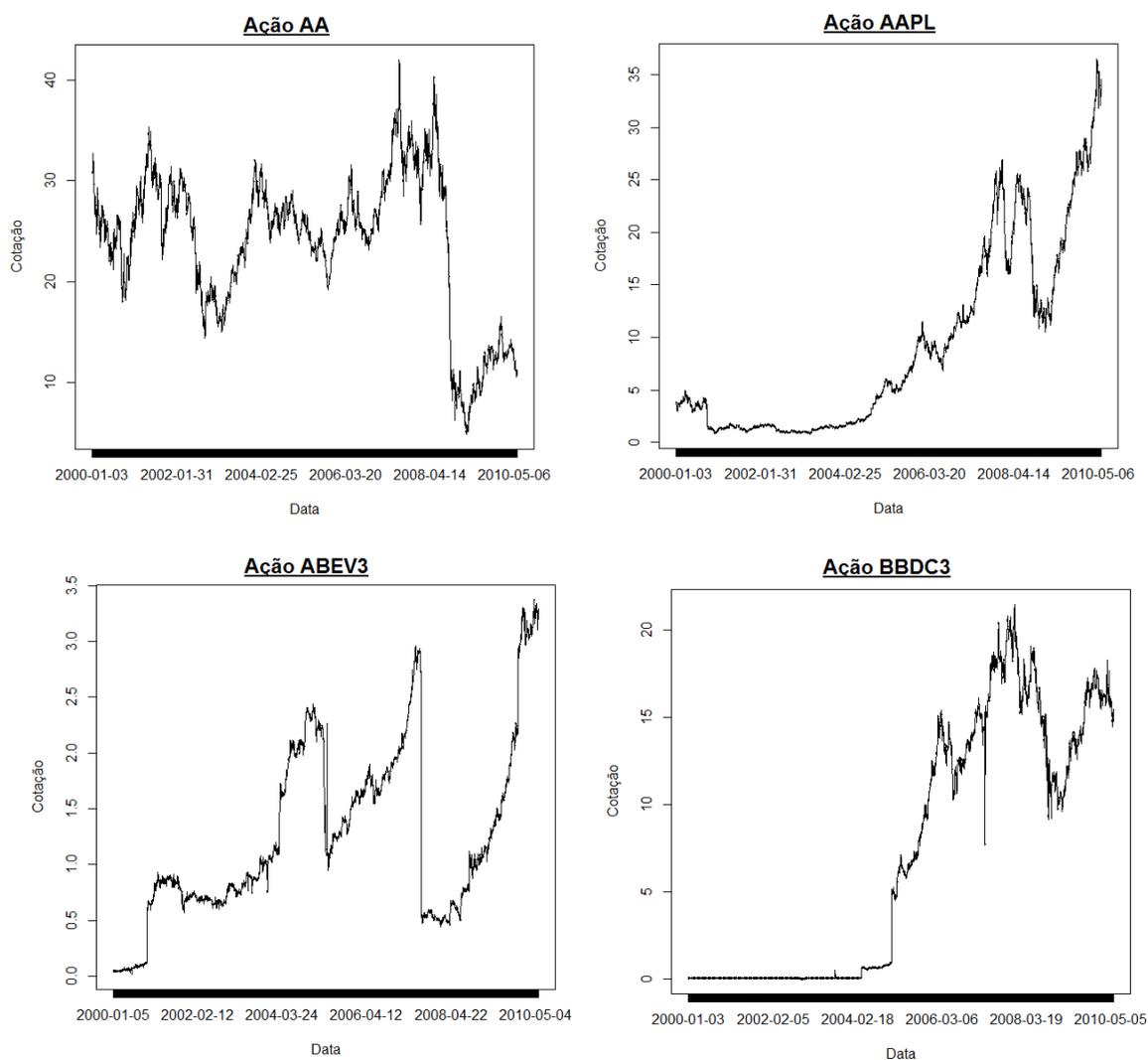


Figura A.1: Série temporal da ação AA, AAPL, ABEV3 e BBDC3.

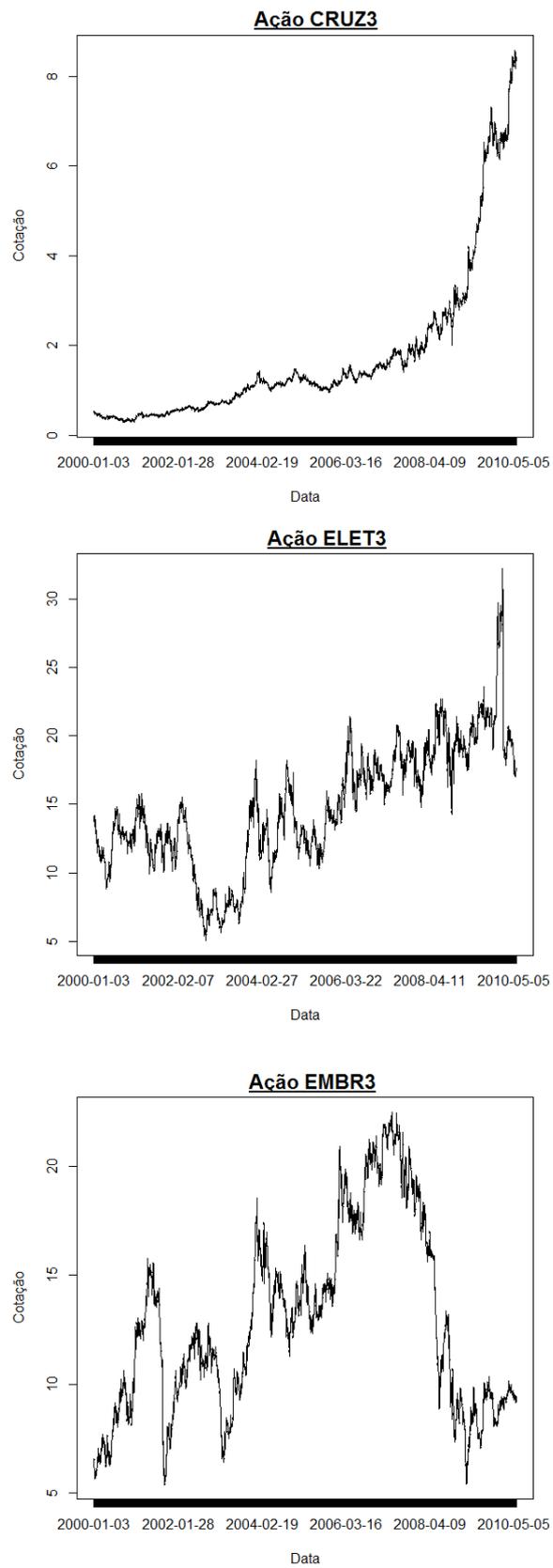


Figura A.2: Série temporal das ações CRUZ3, ELET3, EMBR3.

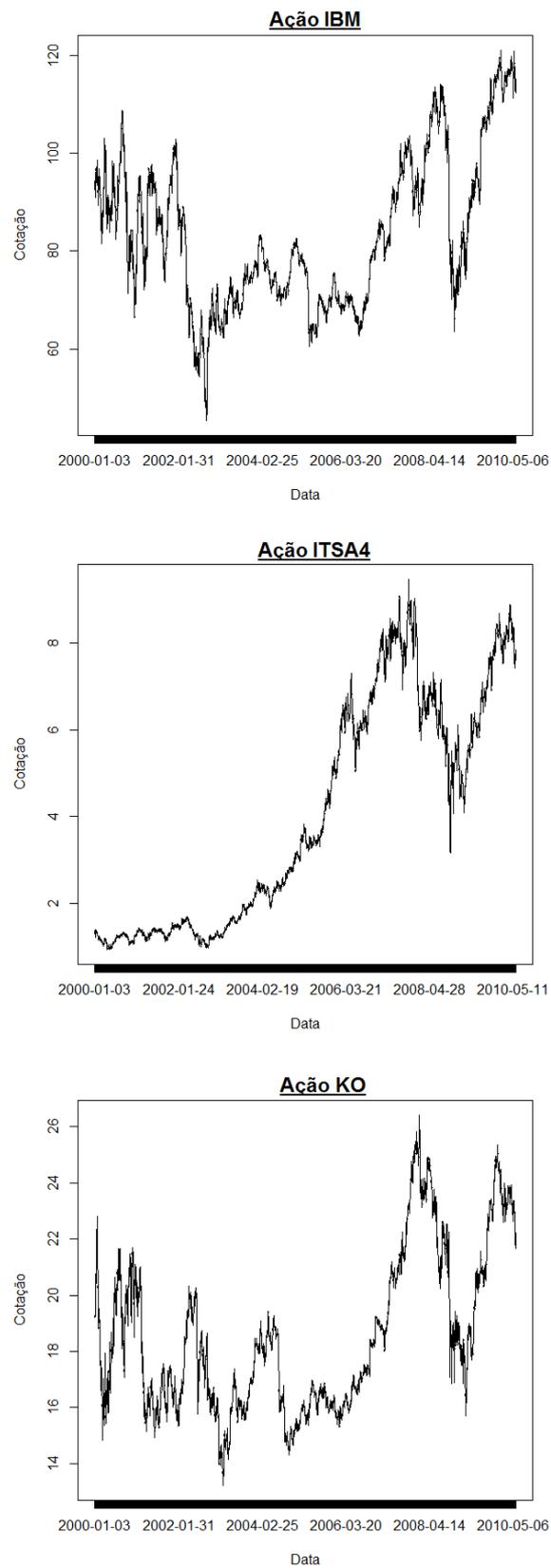


Figura A.3: Série temporal das ações IBM, ITSA4, KO.

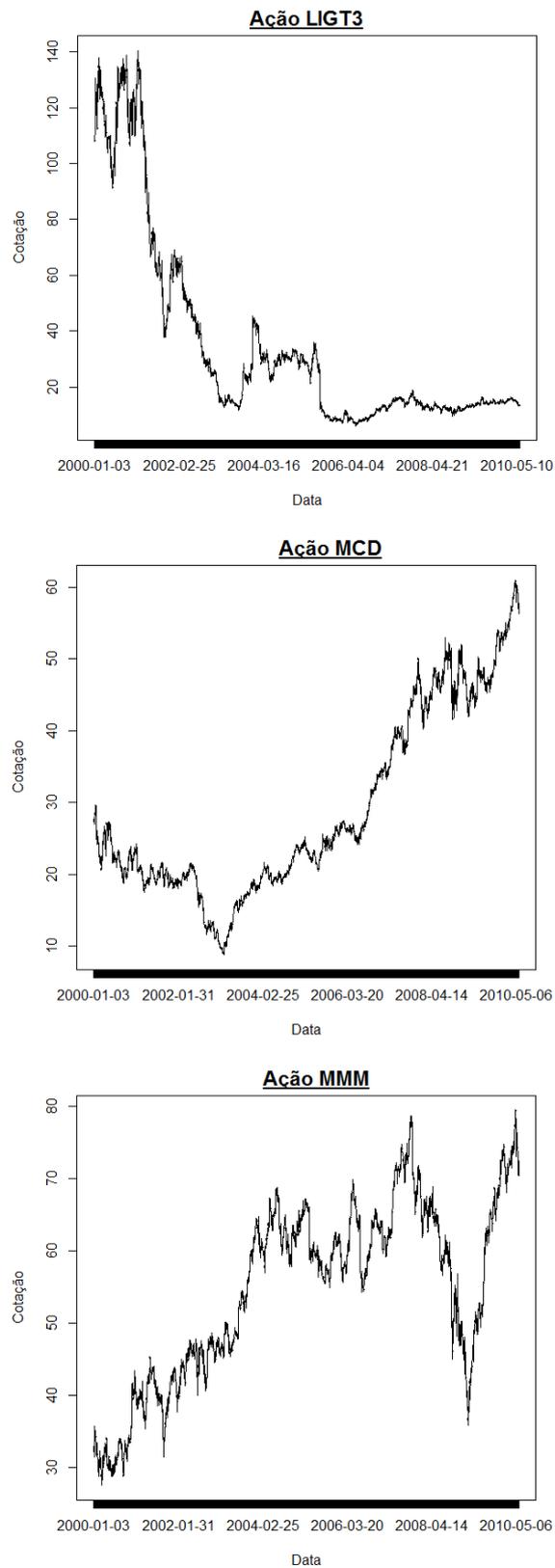


Figura A.4: Série temporal das ações LIGT3, MCD, MMM.

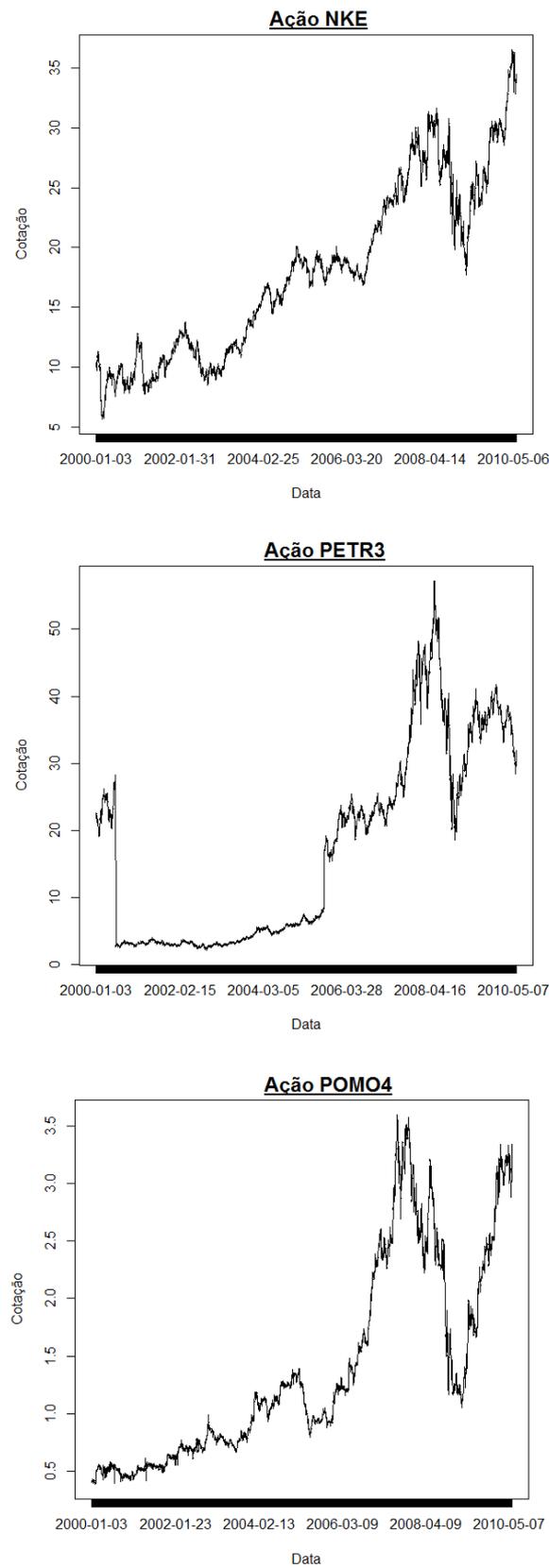


Figura A.5: Série temporal das ações NKE, PETR3, POMO4.

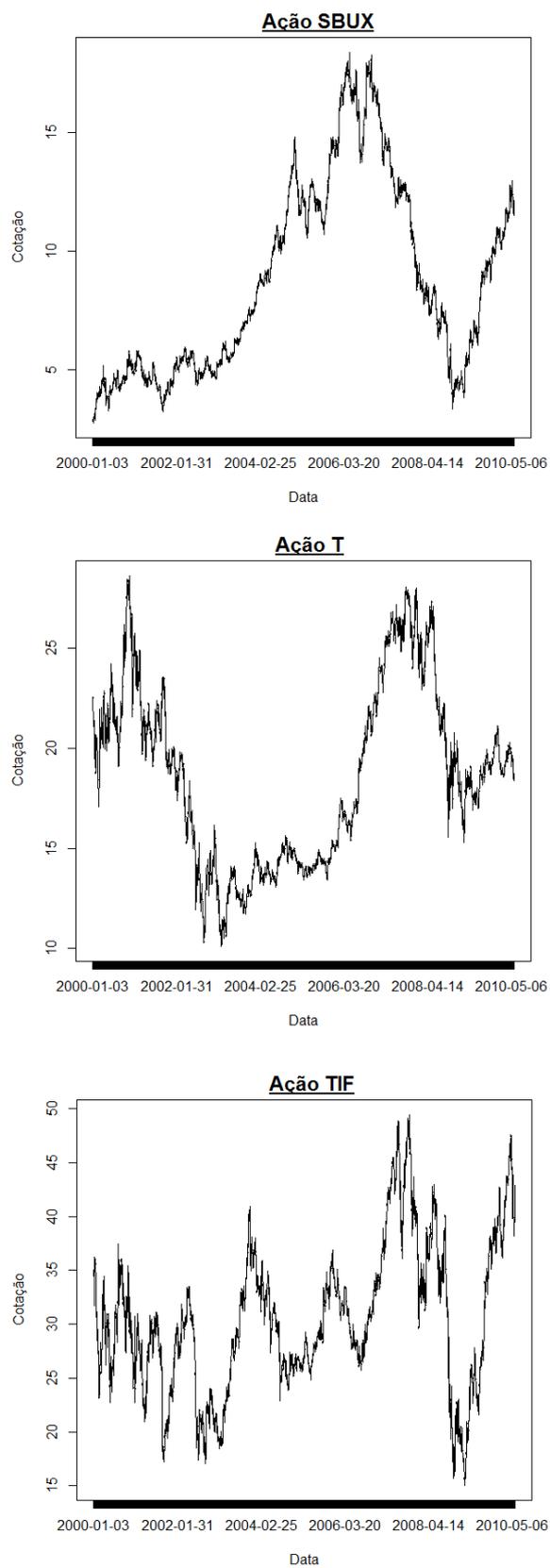


Figura A.6: Série temporal das ações SBUX, T, TIF.

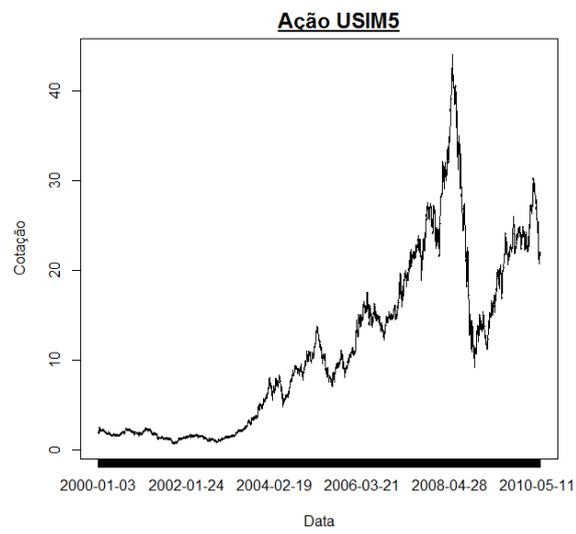


Figura A.7: Série temporal da ação USIM5.