



REDES NEURAIS ARTIFICIAIS NA PREDIÇÃO DAS PRINCIPAIS SÉRIES
DO ÍNDICE IBOVESPA E SUAS APLICAÇÕES EM SISTEMAS
AUTOMATIZADOS DE NEGOCIAÇÃO

Igor Ramalho Pommeranzenbaum

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Luiz Pereira Calôba

Rio de Janeiro
Março de 2014

REDES NEURAIS ARTIFICIAIS NA PREDIÇÃO DAS PRINCIPAIS SÉRIES
DO ÍNDICE IBOVESPA E SUAS APLICAÇÕES EM SISTEMAS
AUTOMATIZADOS DE NEGOCIAÇÃO

Igor Ramalho Pommeranzenbaum

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
ELÉTRICA.

Examinada por:

Prof. Luiz Pereira Calôba, Dr.Ing.

Prof. José Manoel de Seixas, D.Sc.

Prof.^a Marley Maria Bernardes Rebuzzi Vellasco, Ph.D.

Prof.^a Beatriz de Souza Leite Pires de Lima, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2014

Ramalho Pommeranzenbaum, Igor

Redes Neurais Artificiais na Predição das Principais Séries do Índice Ibovespa e suas Aplicações em Sistemas Automatizados de Negociação/Igor Ramalho Pommeranzenbaum. – Rio de Janeiro: UFRJ/COPPE, 2014.

XII, 114 p.: il.; 29, 7cm.

Orientador: Luiz Pereira Calôba

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2014.

Referências Bibliográficas: p. 109 – 114.

1. Redes Neurais Artificiais. 2. Predição de Séries Temporais Financeiras. 3. Sistemas Automatizados de Negociação. I. Pereira Calôba, Luiz. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

Agradecimentos

Gostaria de agradecer a todos que me apoiaram neste projeto.

Aos meus pais, José (*in memorium*) e Celeste, e minha irmã, Luciana.

À minha namorada e meu amor, Paula, seus pais e sua família, que deram fundamental apoio.

Ao meu orientador, pelos direcionamentos de pesquisa.

Aos membros da banca, por terem aceitado o convite.

Aos professores do Mestrado, que foram excelentes na transmissão da informação.

Enfim, obrigado!

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

REDES NEURAI ARTIFICIAIS NA PREDIÇÃO DAS PRINCIPAIS SÉRIES
DO ÍNDICE IBOVESPA E SUAS APLICAÇÕES EM SISTEMAS
AUTOMATIZADOS DE NEGOCIAÇÃO

Igor Ramalho Pommeranzenbaum

Março/2014

Orientador: Luiz Pereira Calôba

Programa: Engenharia Elétrica

A aplicação de técnicas de predição em séries temporais financeiras é um assunto de constante e elevado interesse na comunidade, tanto para investidores quanto para pesquisadores. É uma área desafiadora em relação à complexidade de problemas e que pode gerar altos lucros financeiros para empresas envolvidas com análise de ativos e negociação automática em bolsas de valores. Apresenta-se, nesta dissertação, um modelo de predição de valores futuros sobre as principais séries de preços do índice Ibovespa - a saber: fechamento (*close*), máxima do dia (*high*), mínima do dia (*low*) e a ordem que a máxima e a mínima ocorrem (*order*) - utilizando Redes Neurais Artificiais (RNAs). Para inclusão no modelo, foram obtidas séries temporais de indicadores mundiais, que foram submetidas a métodos clássicos de pré-processamento de séries temporais, para então serem selecionadas como entradas das RNAs por um critério de correlação cruzada. Por fim, foi criado um ambiente de simulação para efetuar a aplicação dos resultados obtidos pelas RNAs utilizando diferentes estratégias de mercado, mais conhecido como *Trading System*. A saída deste simulador são indicadores numéricos que retratam como seria o resultado de uma hipotética aplicação financeira no mercado do mundo real.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

ARTIFICIAL NEURAL NETWORKS IN THE MAIN IBOVESPA INDEX
SERIES PREDICTION AND ITS APPLICATION ON AUTOMATED
TRADING SYSTEMS

Igor Ramalho Pommeranzenbaum

March/2014

Advisor: Luiz Pereira Calôba

Department: Electrical Engineering

The application of techniques in financial time series prediction is a subject of constant and high interest in the community, both for investors and researchers. It is a challenging area in relation to the complexity of problems and can generate high financial returns for companies involved with asset analysis and automated trading on stock exchanges. In this work, we present a model for predicting future values on the main series of Ibovespa index - namely close, high, low and order, which means the order that the high and low occur - using Artificial Neural Networks (ANNs). For inclusion in the model, various time series of global market indexes were obtained, being submitted to classical time series pre-processing methods, so that could be selected as inputs to the ANNs by a cross-correlation criterion. Finally, a simulation environment was created to perform the application of the results obtained by the ANNs using different market strategies, better known as Trading System. The output of this simulator is a set of numerical indicators that show how would be the results for a hypothetical financial market application in the real world.

Sumário

Lista de Figuras	x
Lista de Tabelas	xii
1 Introdução	1
1.1 Visão Geral	1
1.2 Objetivos	3
1.3 Contribuições da pesquisa	5
1.4 Organização do trabalho	5
2 Redes Neurais Artificiais e Predição de Séries Temporais	7
2.1 Fundamentos de Redes Neurais Artificiais	7
2.1.1 Inspiração na estrutura do cérebro humano	8
2.1.2 Vantagens e Desvantagens	9
2.1.3 Principais aspectos estruturais	10
2.1.4 Função de transferência	11
2.1.5 Arquiteturas	12
2.1.6 Parâmetros de treinamento	14
2.1.7 Medidas de erro	15
2.1.8 Teste de performance <i>out-of-sample</i>	17
2.1.9 Outros métodos de IC	18
2.2 Séries Temporais	19
2.2.1 Predição de séries temporais financeiras	20
2.3 Decomposição de Séries Temporais Financeiras	21
2.3.1 Transformação em série de retorno	22
2.3.2 Normalização dos dados	22
2.3.3 Extração de tendência	23
2.3.4 Extração de sazonalidades e ciclos senoidais	23
2.3.5 Atenuação de intrusos (<i>clamping</i>)	24
2.3.6 Estatísticas	24
2.4 Estado da Arte	25

3	Mercado Financeiro	28
3.1	Estrutura do mercado financeiro	28
3.1.1	Histórico da BM&FBOVESPA	29
3.1.2	Índice Ibovespa	29
3.2	Séries temporais financeiras	30
3.2.1	Séries Temporais consideradas neste trabalho	32
3.2.2	Sobre a previsibilidade no mercado financeiro	47
3.3	Formas de Análise em Investimentos	49
3.3.1	Análise Fundamentalista	50
3.3.2	Análise Técnica	51
3.3.3	Gráficos das séries temporais de Análise Técnica	55
3.4	Modelos quantitativos	58
3.5	Sistemas de negociação automática	59
3.5.1	Testes em Trading Systems	61
4	Metodologia	63
4.1	Método criado	63
4.2	Aplicação de janelamento	65
4.3	Obtenção dos dados	66
4.3.1	Tratamento de feriados/dados faltantes	68
4.3.2	Horários de bolsas	68
4.4	Módulo de Pré-Processamento das séries	70
4.4.1	Transformação em série de retornos	70
4.4.2	Normalização	71
4.4.3	Tratamentos de séries temporais	71
4.4.4	Atenuação de intrusos	72
4.4.5	Estatísticas sobre os dados	73
4.5	Critério de seleção de entradas significativas	74
4.6	Módulo de Redes Neurais Artificiais	75
4.6.1	Varição de parâmetros	76
4.6.2	Avaliação do desempenho da RNA	77
4.6.3	Análise da qualidade das predições por faixas de retorno	78
4.6.4	Estudo Comparativo	79
4.7	Módulo de Trading System	80
4.7.1	Regras do Trading System	80
4.7.2	Estratégias implementadas	83
4.7.3	Processamento de Carteira e Indicadores	85

5	Resultados	87
5.1	Experimentos	87
5.2	Resultados - Experimento geral	87
5.2.1	Resultados para as séries alvo	89
5.2.2	Resultados para Ibovespa order - Classificação	93
5.2.3	Resultados de Trading Systems	95
5.2.4	Tempo de execução total	98
5.3	Resultados específicos dos blocos	98
6	Conclusões	106
6.1	Conclusão	106
6.2	Trabalhos Futuros	108
	Referências Bibliográficas	109

Lista de Figuras

2.1	Neurônio artificial	10
2.2	Rede MLP	11
2.3	Gráficos das funções de ativação exemplificadas	12
2.4	Esquema de janelamento clássico (por Kaastra [30])	18
3.1	Gráficos com as séries diárias de <i>close</i> , <i>open</i> , <i>high</i> e <i>low</i> do Ibovespa	33
3.2	Gráfico da série diária do <i>order</i>	33
3.3	Gráfico com a série de Volume do Ibovespa	34
3.4	Gráficos das séries diárias de Feriado nos EUA e Vésperas de Feriado no Brasil	34
3.5	Gráficos das séries de Variação diária e variação máxima do Ibovespa	35
3.6	Gráficos das séries diárias dos índices da Argentina e do Canadá	36
3.7	Gráficos das séries diárias dos índices de Chile e México	37
3.8	Gráficos das séries diárias dos índices americanos: DJIA e S&P 500	38
3.9	Gráficos das séries diárias dos índices da China e de Hong Kong	39
3.10	Gráficos das séries diárias dos índices da Índia e do Japão	40
3.11	Gráficos das séries diárias dos índices da Coreia do Sul e da Rússia	41
3.12	Gráficos das séries diárias dos índices de França e Alemanha	42
3.13	Gráficos das séries diárias dos índices de Itália e Inglaterra	43
3.14	Gráficos das séries diárias do CDI e do Futuro de Ibovespa	44
3.15	Gráficos das séries diárias de Dólar PTAX e Euro	45
3.16	Gráficos das séries diárias de FGV100 e IDIV	46
3.17	Gráficos das séries diárias dos índices IMOB e SMLL	47
3.18	Gráficos das séries de Médias móveis	56
3.19	Gráficos das séries de MACD e IFR	56
3.20	Gráficos das séries de IVR e Bollinger Bands	57
3.21	Gráficos das séries de Momentum e OBV	57
3.22	Gráficos das séries de Volatilidade Anualizada	58
4.1	Esquema dos processos do sistema desenvolvido	64
4.2	Esquema de janelamento utilizado pelo sistema	65

4.3	Exemplo de espectrograma e autocorrelação para a série Ibovespa <i>close</i>	72
4.4	Exemplo de atenuação de outliers	73
4.5	Exemplo de faixas de acurácia	79
5.1	Resultados finais - Modelo RNA - Ibovespa <i>close</i>	89
5.2	Resultados finais - Modelo Linear - Ibovespa <i>close</i>	90
5.3	Resultados finais - Modelo RNA - Ibovespa <i>high</i>	90
5.4	Resultados finais - Modelo Linear - Ibovespa <i>high</i>	91
5.5	Resultados finais - Modelo RNA - Ibovespa <i>low</i>	91
5.6	Resultados finais - Modelo Linear - Ibovespa <i>low</i>	92
5.7	Resultados finais - Modelo RNA - Ibovespa <i>order</i>	92
5.8	Resultados finais - Modelo Linear - Ibovespa <i>order</i>	93
5.9	Histograma de classificação - Modelo RNA - Ibovespa <i>order</i>	94
5.10	Histograma de classificação - Modelo Linear - Ibovespa <i>order</i>	95
5.11	Gráfico de <i>backtesting</i> , com os pontos de compra (triângulos verdes) e de venda (triângulos vermelhos) da melhor estratégia	97
5.12	Gráfico comparativo da evolução dos Preços e do PL da carteira	97
5.13	Exemplo de transformação em série de retorno e normalização, com histogramas - Ibovespa <i>close</i>	99
5.14	Exemplo de Extração de Tendência - Ibovespa <i>close</i>	100
5.15	Exemplo de Espectrograma e Autocorrelação - Ibovespa <i>close</i>	100
5.16	Exemplo de atenuação de valores extremos - Ibovespa <i>close</i>	101
5.17	Exemplo de Correlação significativa - Canada S&P/TSX Index em relação ao Ibovespa <i>close</i> - Lags escolhidos: 1 e 5	102
5.18	Exemplo de gráfico de acompanhamento dos erros de treinamento, validação e teste <i>versus</i> épocas de treinamento - Ibovespa <i>high</i>	105
5.19	Exemplo de gráfico de comparação entre valores reais (linha azul) e estimados (linha vermelha), para os conjuntos de treinamento e teste - Ibovespa <i>high</i>	105

Lista de Tabelas

3.1	Composição atual do índice Ibovespa	31
4.1	Tabela de horários de funcionamento das bolsas mundiais	69
4.2	Tabela com os possíveis valores para cada parâmetro mutável no esquema de variação	77
5.1	Tabela com o número de variáveis significativas escolhidas para cada bloco de treinamento	88
5.2	Tabela com os resultados obtidos para os modelos, comparando a utilização de modelos com RNA e com RNA Linear	88
5.3	Resultados finais de classificação - Ibovespa <i>order</i>	94
5.4	Tabela de resultados da aplicação do <i>Trading System</i> , com comparação entre as estratégias	96
5.5	Tabela com os trades realizados pela melhor estratégia, no período analisado	96
5.6	Exemplo de tabela com as estatísticas, variáveis <i>close</i> , <i>high</i> , <i>low</i> e <i>order</i> do Ibovespa	102
5.7	Variáveis e <i>lags</i> selecionados no bloco nº 29 para as RNAs das séries alvo: <i>Close</i> e <i>High</i>	103
5.8	Variáveis e <i>lags</i> selecionados no bloco nº 29 para as RNAs das séries alvo: <i>Low</i> e <i>Order</i>	104

Capítulo 1

Introdução

1.1 Visão Geral

A aplicação de técnicas de predição em séries temporais financeiras é um assunto de constante e elevado interesse na comunidade científica, tanto para investidores quanto para pesquisadores. É uma área desafiadora em relação à complexidade de problemas e que pode gerar altos lucros financeiros para empresas envolvidas com análise de ativos e negociação automática em bolsas de valores.

Para a resolução de tal problema, deve ser estimado um modelo que melhor aproxima a relação entre as entradas e a saída desejada (denominada "alvo"). Em seguida, o modelo é extrapolado para estimar o valor futuro da série alvo. Há pesquisas que almejam realizar a predição para mais de um período, ou seja, para $t + 2, \dots, t + N$, mas o nível de erro acumulado encontrado neste tipo de modelo normalmente é proibitivo para sua utilização.

Dada a natureza não-linear e de incerteza das séries temporais financeiras, nota-se que onde havia uma preferência por modelos estatísticos, recentemente há um maior interesse em ferramentas oriundas da área de Inteligência Artificial (IA) [5]. Uma de suas principais subáreas, a Inteligência Computacional (IC) - ou *soft computing*, como referenciam alguns trabalhos - tem sido a opção preferida dos autores, englobando principalmente Redes Neurais Artificiais (RNA), Máquinas de Vetor Suporte (em inglês, *Support Vector Machine* - SVM), Algoritmos Genéticos (AG) e Lógica Difusa.

A opção por RNAs tem sido a mais efetuada, motivada principalmente por suas inerentes características. De forma resumida, uma RNA é um aproximador universal que pode mapear qualquer função não-linear, caso apresente múltiplas camadas com neurônios não-lineares [19]. Logo, é uma ótima ferramenta para busca por padrões não-lineares, além de apresentar flexibilidade para lidar com irregularidades. Ainda, é capaz de lidar com dados incompletos ou insuficientes, se mostrando, assim, uma

poderosa metodologia para reconhecimento de padrões, classificação e previsão [34] [28].

A temática deste trabalho lida, portanto, com modelos de predição para um passo adiante (em se tratando de séries temporais econômicas diárias, a notação utilizada é $D + 1$) sobre as principais séries de preços do índice Ibovespa - a saber: fechamento (*close*), máxima do dia (*high*), mínima do dia (*low*) e a ordem que a máxima e a mínima ocorrem (*order*) -, ao contrário da maioria das pesquisas, que buscam apenas tendência ou uma ação a tomar no mercado (compra ou venda) [24].

Analisou-se a influência de diversas séries de entrada, referentes a diversos índices mundiais e indicadores econômicos. As possíveis entradas do sistema são obtidas, geralmente, por experiência dos investidores e traders, ou seja, especialistas do mercado.

Para inclusão no modelo, as séries dos indicadores mencionados foram submetidas a alguns métodos clássicos de pré-processamento de séries temporais, para então serem selecionadas como entradas das RNAs por um critério definido. Os resultados das RNAs, ou mais especificamente, os valores estimados das séries alvo em $D + 1$, são então levados em conta para a ativação de estratégias de negociação automáticas. Tais estratégias tem a responsabilidade de emitir sinais de compra e venda para determinados ativos, no mercado financeiro.

Apesar do assunto já ter sido amplamente explorado mundialmente na literatura, possuindo um grande volume de pesquisas desde o fim da década de oitenta [17] [24], há ainda uma enorme gama de possibilidades de aplicação das mais diferentes técnicas. A grande base de discussões encontrada na literatura envolve tanto a parte econômica do processo - com relação ao retorno e ao risco de estratégias -, quanto suas vertentes matemática e estatística - com a sugestão dos mais variados modelos pelos autores.

Tal amplitude de trabalhos se deve à complexidade que o problema apresenta, pela natureza dinâmica e caótica dos mercados mundiais e pela dificuldade em se obter resultados, de fato, lucrativos. Outra explicação possível é o viés notado por Granger [27], em que é prática comum e entendível não publicar um estudo com excelentes resultados e, ao invés disso, lucrar com o método aplicando-o no mercado real.

No mercado financeiro brasileiro, a abordagem científica ainda não tem a mesma representatividade quando comparada a mercados desenvolvidos, como o norte-americano, alguns europeus e alguns asiáticos. O que é surpreendente, visto que a BM&FBOVESPA, a Bolsa de Valores oficial do Brasil, é a décima maior do mundo em capitalização de mercado (US\$ 1,5 trilhão), de acordo com a Forbes [1] e o Brasil é um mercado emergente, que apresenta diversas oportunidades de investimento.

Apesar do baixo número relativo de estudos no Brasil, já é possível perceber al-

guns sinais de avanço. Diversas empresas, em sua maioria gestoras independentes de fundos de investimento, têm investido cada vez mais em pesquisa e desenvolvimento para aperfeiçoar suas estratégias de negociação de ativos. Usualmente, as pesquisas são relacionadas a diversos campos de aplicação [6], destacando-se: busca de padrões de mercado; previsões sobre preços ou volatilidade de ativos, como: ações, opções de ação, índices, futuros e outros indicadores; alocação ótima de investimentos; dentre diversas outras possibilidades.

A ideia principal destas empresas é a criação e aprimoramento de *trading systems* - termo em inglês para sistemas de negociação automatizada -, com elevados níveis de otimização e eficiência, muitas vezes utilizando-os como próprios condutores de fundos quantitativos. Os fundos quantitativos (ou fundos *quant*, como são chamados) são fundos que efetuam as operações e gestão da carteira seguindo estratégias baseadas em modelos matemáticos e estatísticos, derivados de teorias de áreas multidisciplinares, como engenharia financeira e estatística. Um maior detalhamento sobre estes fundos será exibido no capítulo 3 deste trabalho.

Apesar das necessidades dos fundos quantitativos, um dos pontos críticos encontrados em diversos trabalhos pesquisados é a falta de uma melhor investigação sobre os resultados práticos da previsão. Em muitos, os autores se atêm aos (também importantes) resultados matemáticos dos modelos, como erros médios quadráticos, percentuais ou absolutos. Mas nem sempre se pode afirmar que um determinado modelo é lucrativo apenas olhando para estas variáveis, tornando-se necessário investigar outros indicadores.

Portanto, foi criado um ambiente de simulação para efetuar a aplicação dos resultados obtidos pelas RNAs utilizando diferentes estratégias de mercado, processo tal que é comumente denominado *backtesting*. A saída deste simulador é composta por diversos indicadores numéricos, que exibem como seria o resultado de uma hipotética aplicação do *trading system* no mercado, com dados reais.

1.2 Objetivos

Apesar das vastas possibilidades de aplicação, a criação de métodos precisos para previsão de preços - capazes de serem traduzidas em lucros no mundo real - ainda se apresenta como um grande desafio [34]. Suas principais dificuldades têm relação com a inerente complexidade dos mercados, aliada à não-linearidade e à grande quantidade de incerteza e ruído apresentados pelas entradas [4]. Destas características, se evidenciam as limitações de métodos estatísticos clássicos e a necessidade de utilização de uma técnica mais apurada para as previsões.

Este trabalho, portanto, pretende avaliar o desempenho da aplicação de RNAs na previsão das principais séries de preço do índice Ibovespa, mais especificamente

em dois aspectos: em termos de performance de aproximação e generalização dos modelos de RNA criados para cada uma das séries alvo; e em termos de performance de lucratividade simulada no ambiente de *Trading System* proposto neste trabalho.

Em outras palavras, pretende-se: avaliar o desempenho de uma estratégia automatizada de negociação, baseada em Redes Neurais Artificiais, que realiza previsões sobre algumas das principais séries temporais do índice Ibovespa, a relembrar: *close*, *high*, *low* e *order*.

Com isso, tenta-se estabelecer uma estratégia de operação que se aproveite das previsões para guiar as ordens de compra e venda. Ainda, pretende-se avaliar o desempenho financeiro da estratégia sugerida, realizando uma simulação de possíveis operações no mercado com dados passados por meio do *backtesting*.

Uma justificativa pelo uso do índice Ibovespa é sua representatividade sobre a situação atual do mercado brasileiro. Além disso, foi escolhida a série que representa o contrato futuro de Ibovespa, rolado durante todo o período¹, para as operações intradiárias do *Trading System*, um ativo operável e com ótimos índices de liquidez e volume de negócios diários.

É também um objetivo estudar a influência dos diversos índices mundiais e indicadores econômicos - utilizados como entradas dos modelos - sobre a variação do índice Ibovespa. Um resultado interessante é dado pela análise, em cada período do janelamento, sobre as variáveis significativas e seus respectivos atrasos selecionados.

A fim de testar o modelo proposto, foram obtidos indicadores de desempenho usuais de RNAs, como: raiz do erro médio quadrático (RMSE), erro médio absoluto percentual (MAPE) e coeficiente de regressão, comparando-as com os resultados de outros trabalhos na área.

De forma a investigar mais profundamente os resultados obtidos pelos modelos, foram aplicados os valores de saída do modelo em um sistema de negociação, de forma que fossem obtidas estatísticas mais úteis para investidores, como retorno anualizado e máximo *drawdown* da estratégia. O *drawdown* é um indicador que representa o valor máximo de perda, dado o topo histórico de valor de uma determinada carteira de investimento.

Por fim, algumas comparações foram realizadas. No caso das RNAs, a comparação foi em relação a um modelo linear, implementado por meio de RNAs com apenas uma camada e com um neurônio linear. Já para o *Trading System*, foram comparadas três estratégias de negociação em três diferentes unidades de janela de análise, de forma a escolher a estratégia com melhores resultados.

¹A série rolada é composta pela agregação das séries dos contratos futuros com vencimentos mais próximos, agregados por um fator de rolamento, definido pela BM&FBovespa.

1.3 Contribuições da pesquisa

Com o desenvolvimento desta pesquisa, diferentes questões sobre a aplicabilidade de RNAs para o problema de predição de séries temporais financeiras podem ser avaliadas e debatidas. Algumas, mostradas a seguir, podem ser destacadas:

- Se o mercado é completamente eficiente como afirma a Hipótese dos Mercados Eficientes (EMH) [39], logo, não podendo ser previsível, ou se apresenta alguma possibilidade de predição, refutando a ideia da EMH.
- Se RNAs podem ser consideradas um bom método de predição sobre os indicadores do mercado, de forma que estes valores, aplicados a uma estratégia consistente, gerem lucros para o operador.
- Que variáveis possuem a maior influência na predição de cada série alvo do índice Ibovespa.

Os principais ganhos percebidos nesta pesquisa compreendem: a análise de diferentes abordagens de RNAs para problemas de regressão e classificação; a utilização de um modelo adaptativo que se renova, por meio do janelamento e o reprocessamento das séries, a cada período de aplicação; e a simulação operacional muito próxima à realidade do mercado, dada a operação do ativo real (contrato de Futuro de Ibovespa - IND\$) e a implementação de restrições e contabilidade de custos transacionais.

Vale lembrar que o intuito deste trabalho não é oferecer um método de negociação definitivo, que garanta rentabilidades ou que mitigue o risco. Estratégias sobre ativos de renda variável possuem rentabilidades variáveis e sempre apresentam risco.

Assim como afirma uma máxima da Economia, "rentabilidades passadas não são garantia de ganhos futuros", não é garantido que o desempenho deste sistema será o mesmo descrito aqui durante outros períodos que não o utilizado para testes nesta dissertação.

1.4 Organização do trabalho

Além desta introdução, o presente texto contém mais cinco capítulos, organizados da forma descrita a seguir.

O capítulo 2 discorre sobre o estado da arte de aplicações de Redes Neurais e de sistemas de previsão de séries temporais. Uma revisão bibliográfica é apresentada, bem como os principais trabalhos que envolvem finanças e economia. Além disso, uma breve e importante introdução teórica ao assunto é exibida, de forma que os conceitos utilizados neste trabalho sejam explicados.

Já o capítulo 3 traz uma melhor explanação sobre os conceitos abordados no mercado financeiro, discorrendo sobre os ativos, as empresas participantes e as tecnologias mais utilizadas. Ainda neste, é realizada a contextualização da teoria exibida no capítulo 2 com os conceitos apresentados no capítulo 3, discorrendo sobre a previsibilidade dos mercados financeiros.

Por sua vez, os capítulos 4 e 5 representam o cerne prático deste trabalho. No capítulo 4 é demonstrado o passo a passo da execução dos pré-processamentos, bem como os parâmetros utilizados em cada um deles. No capítulo 5 são explicados cada um dos resultados obtidos, bem como comentada a comparação entre as técnicas investigadas.

Finalizando a parte principal, o capítulo 6 traz a conclusão sobre os estudos realizados, além das possibilidades de aplicação prática dos métodos aqui demonstrados. Como esta é uma área multidisciplinar e muito abrangente, as principais ideias não englobadas no escopo desta pesquisa são sugeridas ao fim de tal capítulo.

Capítulo 2

Redes Neurais Artificiais e Predição de Séries Temporais

Este capítulo será dedicado à descrição dos métodos computacionais e teorias que foram estudados nesta dissertação. Para melhor compreensão, as informações foram divididas em duas partes, que são: redes neurais artificiais (RNA) e séries temporais. Para cada um delas, serão resumidas as bases teóricas e as principais referências de trabalhos na literatura.

2.1 Fundamentos de Redes Neurais Artificiais

O esforço de pesquisa realizado na área de aprendizado de máquina e IC conseguiu produzir um rico conjunto de ferramentas e algoritmos matemáticos, baseado nos trabalhos vistos nas últimas décadas, além de ter dado origem a uma série de aplicações comerciais e científicas. De acordo com [7], a meta de longo prazo da área é produzir métodos que permitam que agentes inteligentes sejam capazes de aprender comportamentos complexos com uma mínima intervenção humana e conhecimento prévio. Exemplos de tais comportamentos complexos são demandados em diversas áreas de aplicações, como percepção visual, percepção auditiva e processamento de linguagem natural.

Uma das principais vertentes de paradigmas de IC, as RNAs são o principal objeto de estudo desta dissertação. RNAs são sistemas compostos por elementos de processamento operando de forma paralela, chamados neurônios, cuja função é determinada pela estrutura de rede, ou melhor, pelos pesos de conexão entre os elementos de processamento, chamados pesos sinápticos, e pelas funções aplicadas por estes elementos.

Ao contrário de outros métodos baseados em modelos tradicionais, RNAs são métodos auto-adaptativos com aprendizagem baseada nos dados experimentais, uti-

lizando assim o conhecimento extraído da experiência de dados sem que existam suposições *a priori* sobre os modelos estudados. Eles aprendem com exemplos e buscam capturar relações lineares ou não-lineares - dependendo da arquitetura utilizada - que descrevam bem o conhecimento desejado. Além disso, apresenta habilidade em lidar com dados ruidosos, incompletos ou imprecisos, o que torna a sua aplicação bastante eficiente em uma enorme gama de problemas.

Na prática, RNAs são especialmente úteis para problemas de classificação e mapeamento e aproximação de funções que sejam tolerantes com imprecisão e que possuam suficiente quantidade de dados disponíveis para treinamento, mas para os quais não há leis de formação ou modelos que possam ser aplicados diretamente, como por exemplo, as provenientes de sistemas especialistas.

Uma arquitetura de RNA com apenas uma camada intermediária de entrada é suficiente para realizar qualquer mapeamento L2, o que envolve a totalidade dos mapeamentos de interesse em nosso caso [28] [12]. Apesar disso, é importante ressaltar que não existe algoritmo para treinamento de RNAs que possa inferir qualquer informação que não esteja contida nos dados de treinamento [7].

RNAs podem ser aplicadas nos mais diversos campos de pesquisa para modelar um determinado problema. As aplicações típicas podem ser encontradas em: diagnósticos médicos, na análise de dados econômicos e financeiros, no controle de processos de produção, em robótica, dentre outras áreas. Atualmente, RNAs estão sendo utilizadas com um enfoque mais comercial, para uma grande variedade de tarefas em diversas áreas de comércio, indústria e ciência [28].

2.1.1 Inspiração na estrutura do cérebro humano

O estudo das RNAs é inspirado no estudo biológico do funcionamento do cérebro humano, obviamente, de uma maneira simplificada quando se compara a estrutura artificial à complexa estrutura biológica. A semelhança é evidenciada pelos elementos equivalentes como neurônios e sinapses entre esses neurônios. No caso das RNA, os neurônios são as unidades de processamento e as sinapses são o meio de comunicação para troca de informação entre esses neurônios.

Conceitualmente, o cérebro é uma grande prova de que o processamento paralelo, tolerante a falhas e adaptativo, é, além de possível, muito rápido e poderoso. As RNAs assemelham-se ao funcionamento do cérebro no sentido em que o conhecimento é adquirido do ambiente pela rede a partir do seu processo de aprendizagem e pela existência das forças de conexão entre os neurônios (pesos sinápticos), utilizadas para armazenar o conhecimento adquirido.

2.1.2 Vantagens e Desvantagens

A aplicação de RNAs traz diversas vantagens quando comparada com a aplicação de outros métodos. As informações descritas aqui têm como principal referência o livro de Haykin [28].

RNAs podem ter neurônios lineares ou não-lineares, sendo que uma RNA que possua ao menos um neurônio não-linear é considerada não-linear. Esta propriedade é especialmente útil para problemas não-lineares, como a predição no mercado financeiro, por exemplo.

Por sua vez, RNAs podem se utilizar da aprendizagem supervisionada, onde aprendem através de exemplos. Isso significa que, a partir de um conjunto com comportamentos esperados e as razões que levam a determinados comportamentos, é possível realizar o mapeamento e fazer com que a RNA aprenda o fenômeno.

Ainda, RNAs possuem uma capacidade natural de adaptação dos pesos de seus neurônios de acordo com modificações no ambiente, podendo ser facilmente retreinadas. Cada neurônio é influenciado pela atividade dos outros, formando automaticamente a noção de contexto.

Falando em aplicações em predição de séries temporais, várias características de RNAs tornam-nas interessantes nesta área. A capacidade de abordar o problema em busca de padrões não-lineares também lhes dá grande vantagem. Além disso, as RNAs são capazes de generalizar, mesmo em situações com certo nível de ruído. As redes neurais são capazes de capturar padrões de ordem superior à linear ou estrutura subjacente dentro de uma série temporal, mesmo quando a lei de formação que a rege seja desconhecida ou complexa demais para descrever.

Como principais desvantagens, pode-se citar a impossibilidade de se criar um método "perfeito para todas as aplicações", conforme [54], em que é citado o teorema *No-Free-Lunch*, concluindo que para todo modelo existe uma amostra de dados em que este apresentará resultados ruins.

Uma rede neural geralmente pode consistir de milhares de sinapses que interagem de acordo com a arquitetura escolhida. Logo, uma importante desvantagem de utilização de um modelo de negociação que se baseie em predições de RNAs, por exemplo, é a dificuldade de interpretar os motivos pelos quais um dado vetor de entrada produziu uma determinada saída, gerando decisões com pouca ou nenhuma base de explicação, dada a natureza de "caixa preta" dos modelos neurais. Em outras palavras, se o modelo toma uma decisão errada, é complicado verificar o motivo do erro, ou se o problema apresenta características recorrentes.

Outra desvantagem evidente é a falta de fórmulas ou heurísticas confiáveis para determinação dos parâmetros a serem utilizados pela RNA. Por fim, vale citar que, para determinados problemas, pode ser altamente necessário que se tenha disponível

um grande conjunto de valores para treinar devidamente o modelo e para que este realize boas inferências.

Apesar de muitas características favoráveis à escolha de RNAs, a construção de um modelo de predição que utiliza RNAs não é uma tarefa trivial. Conforme será visto na próxima seção, há diversos fatores e decisões a serem tomadas que podem determinar o sucesso ou falha da implementação do método.

2.1.3 Principais aspectos estruturais

O neurônio artificial é a unidade de processamento básica das redes neurais, sendo uma simplificação inspirada no funcionamento do neurônio biológico. Este modelo básico de neurônio foi concebido no trabalho de McCulloch-Pitts, em 1943 [41] e forma a base para a maioria dos modelos conexionistas desenvolvidos posteriormente [45]. A representação do neurônio é exibida na figura 2.1, abaixo:

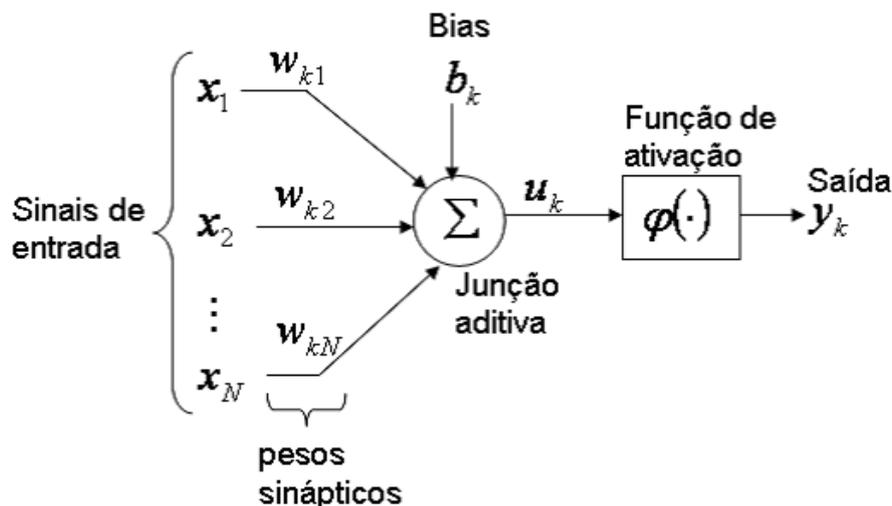


Figura 2.1: Neurônio artificial

Já Rosenblatt mostrou em seu livro "Principles of Neurodynamics", em 1958, o modelo do Perceptron [52]. Nele, os neurônios eram organizados em camadas, uma de entrada e outra de saída, e os pesos das conexões eram adaptados a fim de se atingir a eficiência sináptica. Rosenblatt demonstrou que, se fossem acrescentadas sinapses ajustáveis, as RNAs de McCulloch e Pitts poderiam ser treinadas para classificar padrões em classes linearmente separáveis, convergindo em um número limitado de passos [44].

As redes MLP (*Multi-Layer Perceptron*), portanto, representam uma generalização do Perceptron de camada única, possuindo uma ou mais camadas intermediárias (escondidas) de processamento. Um exemplo de rede MLP totalmente conectada e com duas camadas ocultas é exibida na figura 2.2. Este modelo é o mais utilizado pelos trabalhos encontrados na literatura.

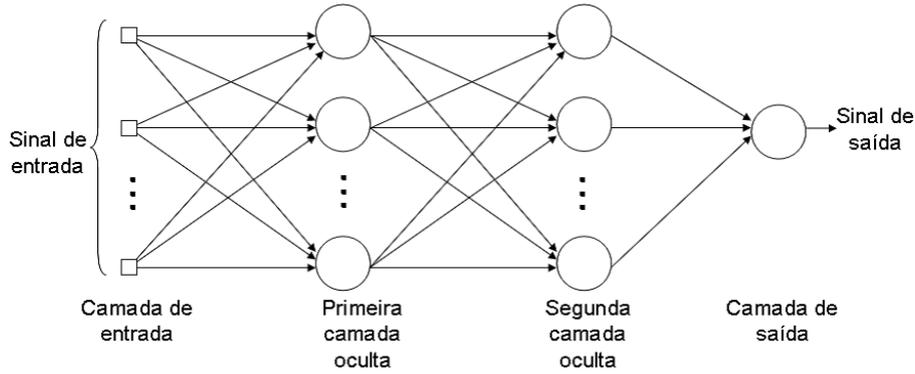


Figura 2.2: Rede MLP

Por fim, as RNAs podem ser classificadas quanto à maneira como os sinais se propagam dentro da de sua estrutura interna, determinando se a rede possui retroalimentação ou não, podendo ser classificadas como alimentadas adiante (*feedforward*) ou recorrentes. Nas redes alimentadas adiante, o fluxo de sinal é apenas em um sentido, como na figura 2.2. Nas redes recorrentes, existe pelo menos um ciclo de retroalimentação, em que o sinal retorna para uma camada anterior [53].

2.1.4 Função de transferência

A função de ativação de um neurônio artificial ($\varphi(\Delta)$) calcula a saída (restringindo a amplitude) do neurônio, em função do valor do potencial de ativação $u(k)$. As funções de ativação $\varphi(\cdot)$ mais comumente utilizadas são: Degrau (Eq. 2.22), Linear (Eq. 2.2), Sigmóide (Eq. 2.3) e Tangente Hiperbólica (Eq. 2.4).

- Função Degrau:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (2.1)$$

- Função Linear:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq \frac{1}{2} \\ v, & \text{se } -\frac{1}{2} < v < \frac{1}{2} \\ -1, & \text{se } v \leq -\frac{1}{2} \end{cases} \quad (2.2)$$

- Função Sigmóide:

$$\varphi(v) = \frac{1}{1 + e^{-av}} \quad (2.3)$$

Onde a é o parâmetro de inclinação da sigmóide. Variando-se o parâmetro a , obtém-se diferentes inclinações da função sigmóide, como pode ser visto na figura 2.3(c).

- Função Tangente Hiperbólica:

$$\varphi(v) = \tanh(v) \quad (2.4)$$

As funções anteriormente descritas são ilustradas na figura 2.3.

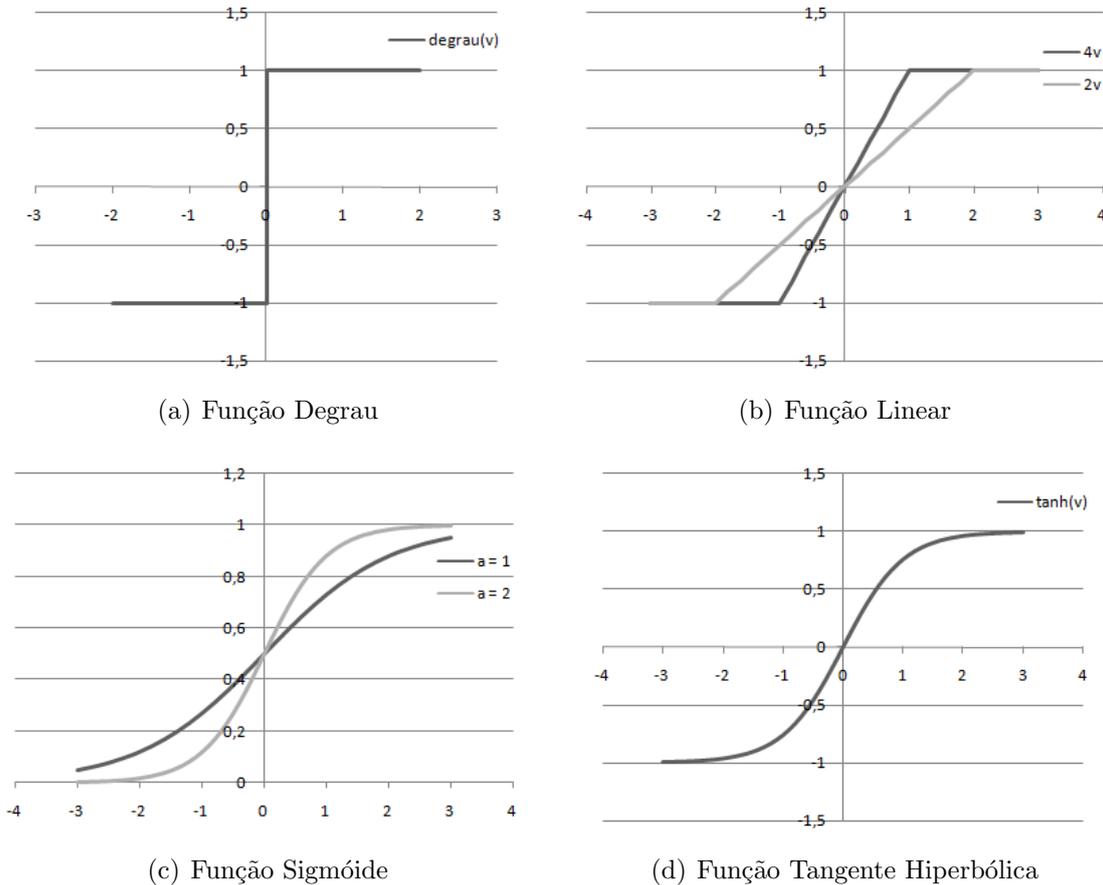


Figura 2.3: Gráficos das funções de ativação exemplificadas

Caso as saídas desejadas sejam contínuas, usualmente faz-se uso de neurônios lineares na saída. Caso contrário, ou seja, as saídas sejam discretas e a RNA deve resultar em uma classificação, utilizam-se neurônios não-lineares. Ainda, em alguns casos, mostra-se útil utilizar um neurônio linear na camada intermediária formada por neurônios não-lineares, de forma que a rede possa expressar alguma relação linear com mais evidência [12].

2.1.5 Arquiteturas

A forma como os neurônios são agrupados é conhecida como arquitetura de RNA. Os neurônios artificiais podem ser organizados e conectados de várias maneiras, resultando em diferentes arquiteturas neurais, com características e aplicações bastante distintas.

Antes da distinção por tipos, há duas grandes divisões que devem ser esclarecidas sobre os diferentes tipos de RNAs. A primeira é dada pela forma como o algoritmo de aprendizado é implementado, sendo as formas mais comuns: os algoritmos supervisionados e os não supervisionados. De forma direta, no aprendizado supervisionado são exibidos os valores corretos para cada par entrada-saída, de forma que a RNA possa ajustar seus pesos e diminuir o erro geral de treinamento. Já no não supervisionado, a RNA realiza uma forma de compressão nos dados, agrupando os valores que apresentam padrões similares.

A segunda distinção diz respeito à forma com que os dados são processados na RNA, podendo ser classificadas em redes diretas (ou *feedforward*) e redes recorrentes. Nas redes *feedforward*, as conexões entre os neurônios não formam ciclos, pois o sentido de propagação é único. Por outro lado, nas redes recorrentes, há ciclos nas conexões, tornando o treinamento mais complexo, mas especializando a rede para certos tipos de aplicação, como os sistemas variantes no tempo.

Atualmente, existem muitos tipos de RNAs, dado que novas arquiteturas são criadas (ou pelo menos, variadas das antigas) a cada momento. O algoritmo de treinamento mais utilizado na literatura é o *Backpropagation*, dada a simplicidade de implementação e provada eficiência. A seguir é exibida uma lista com algumas das arquiteturas (e algoritmos de treinamento) mais conhecidas. Para conceituação e ilustração das arquiteturas citadas a seguir, sugere-se as principais referências da área de RNAs, explanadas com detalhes na seção 2.4 desta dissertação, que discorre sobre o Estado da Arte da literatura.

- Redes MLP
 - Backpropagation
 - Quickprop
 - Resilient Propagation
- Redes RBF
- Redes recorrentes
 - Elman
 - Time-Delay Neural Network
 - Time-Lagged Feedforward Neural Network
- Deep Learning Neural Networks
- Redes não-supervisionadas
 - Kohonen

SOM

ART

- Métodos Híbridos

Modelos Neurofuzzy

Genetic Neural Networks

Com o objetivo de comparação com métodos lineares, RNAs podem ser adaptadas para se comportarem como modelos lineares de regressão. RNAs sem camadas intermediárias e neurônios lineares na camada de saída são basicamente modelos lineares. RNAs com uma camada intermediária e neurônios lineares na camada de saída apresentam grande similaridade com o modelo Projection Pursuit Regression [51].

2.1.6 Parâmetros de treinamento

Para a realização do treinamento de RNAs, uma série de decisões devem ser tomadas de forma que o treinamento convirja para o objetivo, seja a minimização do erro médio quadrático ou qualquer que seja a função a minimizar.

Uma das decisões a serem tomadas é sobre a forma de aprendizado, que pode ser:

- Batelada (*batch*): quando a atualização dos pesos é dada ao fim do processamento de todos os pares entrada-saída. Pode tornar-se lento se o conjunto de treinamento possuir muitos pares;
- Em lotes: quando o conjunto é dividido em lotes para, ao fim do processamento de cada lote, é efetuada a atualização dos pesos;
- Regra delta (treinamento online): quando a atualização dos pesos se dá ao processar cada par entrada-saída apresentado.

Outras importantes decisões sobre o processo de treinamento incluem:

- Número de camadas escondidas;
- Número de neurônios nas camadas escondidas;
- Número de neurônios de saída;
- Valor máximo inicial de distribuição dos pesos sinápticos;
- Taxa de aprendizado.

Estes parâmetros apresentam importante influência nos resultados de predição. Por exemplo, uma RNA com número insuficiente de neurônios na camada escondida pode apresentar uma aproximação ruim da série modelada. Outro exemplo é dado pelo ponto inicial da otimização, dado pelo valor inicial aleatório dos pesos, onde no processo a RNA pode atingir um ponto de mínimo local que apresenta um resultado real impreciso para a predição. Este caso é agravado caso a escolha da taxa de aprendizado (ou passo de treinamento) seja um valor insuficiente para buscar o mínimo global.

Não existe uma fórmula ou heurística definida para a obtenção destes parâmetros, sendo estes normalmente definidos por uma avaliação empírica realizada a partir de vários testes experimentais.

Um fator crítico que pode levar ao sucesso ou falha da aplicação é a coerente definição dos critérios de parada do processo, que podem se dar por:

- Número máximo de épocas atingido;
- Erro mínimo alcançado;
- Um valor de variação mínima para os erros foi atingido.

2.1.7 Medidas de erro

O desempenho das RNAs treinadas deve ser, por fim, avaliado por meio de indicadores de desempenho das mesmas. A função de erro mais comumente avaliada pela literatura [6] é o Erro Médio Quadrático (do inglês *Mean Squared Error*, MSE). Esta e outras mais utilizadas na literatura [28] são apresentadas a seguir, nas equações 2.5, 2.6, 2.7, 2.8 e 2.9:

- Erro Médio Quadrado (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^{i=N} (e_i)^2 \quad (2.5)$$

- Raiz do Erro Médio Quadrático (RMSE):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} (e_i)^2} \quad (2.6)$$

- Erro Médio Percentual Absoluto (MAPE):

$$MAPE = \frac{100}{N} \sum_{i=1}^{i=N} \left| \frac{e_i}{x_i} \right| \quad (2.7)$$

- Correlação: mede a correlação de Pearson entre os valores estimados pela RNA e os valores reais de saída.

$$r = \frac{cov(X, Y)}{\sqrt{var(X)}\sqrt{var(Y)}} \quad (2.8)$$

- Acurácia: define a porcentagem de acertos de sinal para os valores estimados pela RNA, quando comparados aos valores reais da saída.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.9)$$

Onde, dada a matriz de confusão de uma predição (classificação) sobre os valores reais, TP representa os valores verdadeiros positivos, TN são os verdadeiros negativos, FP são os falsos positivos e, por fim, FN representa os falsos negativos.

Ainda, há indicadores de erros específicos extraídos de RNAs de classificação, de forma a obter um melhor detalhamento sobre a análise de seus resultados. Os indicadores apresentados aqui são baseados no trabalho de Kohavi [32], que cita tais indicadores como uma forma de avaliar uma tarefa de classificação binária, sendo todos estes calculados com base nas já conceituadas variáveis da matriz de confusão, a lembrar, TP , TN , FP e FN .

A principal motivação para extração de tais indicadores é a utilização, a seguir, de uma RNA do tipo classificação para predição da série alvo *order*. A listagem das fórmulas para cada medida de classificação se encontra a seguir, nas equações 2.10, 2.11, 2.12, 2.13, 2.14, 2.15 e 2.16, onde a seguir é também explicado o significado de cada indicador.

- Sensibilidade (S)

$$S = \frac{TP}{TP + FN} \quad (2.10)$$

- Especificidade (E)

$$E = \frac{TN}{TN + FP} \quad (2.11)$$

- Valor Preditivo Positivo (VPP)

$$VPP = \frac{TP}{TP + FP} \quad (2.12)$$

- Valor Preditivo Negativo (VPN)

$$VPN = \frac{TN}{TN + FN} \quad (2.13)$$

- Falsos Alarmes (FA)

$$FA = \frac{FP}{TN + FP} = 1 - E \quad (2.14)$$

- Falsas Perdas (FL)

$$FL = \frac{FN}{TP + FN} = 1 - S \quad (2.15)$$

- Erro de Classificação (e_{class})

$$e_{class} = \frac{FP + FN}{TP + TN + FP + FN} \quad (2.16)$$

O indicador S , ou como também é conhecido, Taxa de Verdadeiros Positivos, mede a capacidade do modelo de detectar corretamente os valores positivos. De forma análoga funciona o indicador E , também chamado de Taxa de Verdadeiros Negativos, medindo capacidade de detecção de valores negativos. Por sua vez, a medida VPP indica a probabilidade de um valor classificado como positivo realmente corresponder à sua classificação real. O VPP contrasta com o VPN , no sentido que este último atua de forma similar, mas para valores negativos.

O valor de FA exprime a taxa de erros de detecção de positivos, sendo o complementar do valor de E , similarmente ao indicador FL , que indica a taxa de erros de classificação de valores negativos, sendo complementar à medida S . Por fim, o erro de classificação analisa o erro geral do processo, independente da pertinência dos valores às diferentes classes da matriz de confusão.

2.1.8 Teste de performance *out-of-sample*

É prática comum avaliar o poder de generalização de RNAs por meio da divisão dos dados em três clássicos conjuntos: treinamento, validação e teste (*out-of-sample*¹). De acordo com Kaastra [30], o conjunto de treinamento é o maior deles e é o que as RNAs usarão de fato para aproximar a função objetivo. O conjunto de validação será avaliado a cada execução da RNA para verificar a possibilidade de parada prematura do treinamento e para determinar a melhor época de treinamento da RNA. Por fim, o conjunto de teste seria o responsável por avaliar a generalização e performance do treinamento da RNA em dados *out-of-sample*, usualmente, sobre os dados mais recentes do conjunto.

Este modo de análise é amplamente utilizado na literatura, mas apresenta algumas desvantagens. Uma das principais é que a divisão é comumente dada pela separação do conjuntos em 70/15/15 (ou similares). Isso significa que 70% dos da-

¹Dados *out-of-sample* são aqueles que não são conhecidos (ou amostrados) na fase de treinamento das RNAs.

dos serão separados para treinamento, 15% para validação e os 15% restantes para teste, ficando assim o teste *out-of-sample* muito limitado.

Por tal motivo, Kaastra [30] recomenda utilizar uma abordagem mais rigorosa, chamada janelamento (ou método de janelas deslizantes), baseado em testes *walk-forward*. O esquema deste modelo é exibido na figura 2.4 a seguir. Na figura mostrada aqui, há uma pequena diferença de nomenclatura, que inverte os termos "Testing Set" e "Validation Set", com relação ao esquema de janelamento e divisão dos conjuntos utilizado neste trabalho.

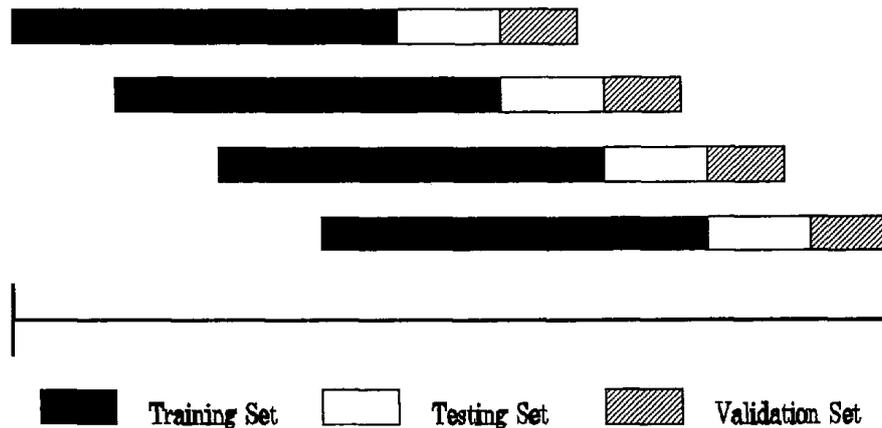


Figura 2.4: Esquema de janelamento clássico (por Kaastra [30])

O janelamento consiste em realizar a divisão treinamento/validação/teste em diversas janelas de tempo, de forma que os conjuntos de teste se sobreponham e a cada bloco, ocorra o deslocamento da janela de análise. O objetivo é simular a aplicação real do sistema, em que a cada determinado número de períodos, ocorre o retreino das RNAs. Apesar de este processo ser computacionalmente mais custoso, há a vantagem das RNAs se adaptarem melhor às mudanças de mercado. Vale lembrar que este modelo foi a principal base para a elaboração do esquema utilizado aqui, que será explicado com maiores detalhes no capítulo 4.

2.1.9 Outros métodos de IC

Os problemas que RNAs solucionam também têm sido abordados por diferentes técnicas, que variam de métodos estatísticos a avançados modelos utilizando Inteligência Computacional (IC) - ou como também são conhecidas, técnicas de *soft computing* -, inclusive com modelos híbridos, misturando tais técnicas.

Os métodos mais comumente encontrados na literatura são:

- Support Vector Machines (SVM) e Support Vector Regression (SVR): [6], [33], [50];

- Algoritmos Genéticos: [6], [33];
- Programação Genética: [23];
- Genetic Network Programming (GNP) - um algoritmo evolutivo baseado em grafos: [14], [15];
- Sistemas Neuro-Fuzzy: [6].

2.2 Séries Temporais

Na literatura [43], uma série temporal é definida como "qualquer conjunto de observações ordenadas no tempo". Em outras palavras [49], as séries temporais pertencem a uma classe de fenômenos cujo processo observacional gera uma sequência de dados distribuídos no tempo, sendo estes dados obtidos em intervalos de tempo durante um período específico. Embora tal definição não restrinja a questão de regularidade dos intervalos, a maioria das séries temporais observadas nos estudos possui observações em intervalos regulares de tempo.

Conceitualmente, uma série temporal com n observações possui a seguinte representação: $X = x_1; x_2; \dots; x_n$, onde X é a série temporal e as variáveis x_1 a x_n são amostras da série no período $t = 1, 2, \dots, n$.

Uma das principais características deste tipo de fenômeno é que as observações vizinhas são dependentes entre si e, em vários casos, é interessante analisar e modelar esta dependência [22]. Enquanto que, em modelos onde o paradigma dos dados é atemporal, por exemplo, a ordem das observações é irrelevante para a resposta final do processo, em análises sobre séries temporais a ordem dos dados é um fator crítico.

Uma série temporal pode ser discreta ou contínua. A série será discreta, se o conjunto de observações for discreto no tempo e será contínua, obviamente, se o conjunto de observações for contínuo. A grande maioria dos métodos de previsão utiliza séries discretas ou discretizadas [53]. A conversão de uma série contínua para discreta pode ser realizada pela medição da série original a cada N pontos em um intervalo definido, ou realizando a acumulação de valores nos subintervalos desejados, sendo sempre necessário verificar se o contexto da série permite tal transformação.

Uma das aplicações mais exploradas por pesquisadores na literatura é a modelagem e previsão de séries temporais financeiras. Exemplos de séries nesta área são listadas aqui:

- Valores diários de preços de fechamento, máximo, mínimo e abertura de um ativo negociado na BM&FBovespa;
- Valores de volatilidade de ativos registrados a cada 15min;

- Volumes de negociação para ativos listados em bolsa;
- Série histórica de índices de inflação ou taxas de juros;
- Série de indicadores de Análise Técnica aplicada a alguma série de preços de um determinado ativo.

Nesta dissertação, são estudadas diversas séries temporais financeiras, que serão descritas com maiores detalhes no capítulo 3. Neste capítulo, nos ateremos à explicação dos pré-processamentos necessários para correta introdução dos valores das séries temporais nas RNAs.

2.2.1 Predição de séries temporais financeiras

Dentro do universo de séries temporais, há um importante segmento com o qual lida este trabalho, que é o das séries temporais financeiras. A modelagem de séries desse grupo tem grande importância para pesquisadores e analistas de mercado, que objetivam efetuar algumas tarefas com mais exatidão e performance, como: a estimação de preços de ativos, de volatilidades, além da obtenção de parâmetros confiáveis para análise de risco e otimização de portfólios de investimento.

O problema de predição de séries temporais pode ser definido como segue: dado um conjunto de exemplos $S = (x_i; y_i); i = 1, \dots, n$ onde para todo $i, f(x_i) = y_i$, retorne uma função g que aproxima f de forma a minimizar a norma do vetor de erro $E = (e_1, \dots, e_n)$. Cada e_i é definido como $e_i = e(g(x_i), y_i)$ onde e é uma função de erro arbitrária [29].

Quando aplicada a uma previsão no mercado financeiro, a definição é análoga, fazendo com que S seja o conjunto de treinamento, com valores x_i representando o vetor de entradas, que pode conter, por exemplo, valores relacionados ao ativo, cotações de outros ativos e/ou indicadores de análise técnica, e, logo, os valores y_i representando o preço a ser previsto.

A aplicação de técnicas de predição no mercado financeiro é uma área desafiadora em relação à complexidade de problemas e que pode gerar altos lucros financeiros para empresas, principalmente bancos de investimento, gestoras independentes de fundos e administradores de fundos.

Modelos clássicos são abordados em abundância na literatura. Um dos mais conhecidos é o modelo desenvolvido pelos professores George E. P. Box e Gwilym M. Jenkins em seu livro [9], o modelo de Box & Jenkins, também conhecido como modelo Auto-Regressivo Integrado de Média Móvel (do inglês *AutoRegressive Integrated Moving Average*, ARIMA).

Também amplamente discutido, o modelo GARCH (*General Auto-Regressive Conditional Heteroscedasticity*) busca a modelagem da variância condicional da

série, em contrapartida à modelagem da média móvel. Esta possui diversas aplicações em modelagem de mercados financeiros, sobretudo na previsão da volatilidade de ativos [21].

Apesar de vastas possibilidades de aplicação, criar métodos de predição de tendências futuras de ativos com precisão, capazes de serem traduzidas em lucros no mundo real, ainda se apresenta como grande desafio [34]. As principais dificuldades do problema de predição de preços têm relação com a inerente complexidade dos mercados, aliada à não-linearidade e à grande quantidade de incerteza e ruído apresentados pelos dados.

Destas características, se evidenciam as limitações de métodos estatísticos clássicos e a necessidade de utilização de uma técnica mais apurada para as predições [5]. A opção por RNAs tem sido a mais efetuada, motivada principalmente por sua capacidade em lidar com características inerentes aos dados: é uma ótima ferramenta para busca por padrões não-lineares, pode lidar facilmente com irregularidades e é capaz de lidar com dados incompletos ou insuficientes.

Apesar de não ser um assunto novo - há registros de pesquisas desde o fim da década de oitenta [17] -, sempre há novos aspectos de pesquisa, dado o padrão dinâmico e caótico dos mercados [4]. As primeiras aplicações de RNAs em predições de preços de ações não foram completamente bem sucedidas, sendo aos poucos descobertas novas metodologias que permitiram às RNAs se mostrarem como boas alternativas para esse tipo de aplicação [17].

Pelo visto, o sucesso na tarefa de predição está relacionado, em sua maior parte, à escolha de variáveis importantes para entrada do modelo [30]. Neste trabalho escolheu-se utilizar índices mundiais e indicadores do mercado como possíveis entradas dos modelos, por experiência de analistas financeiros e também de outros trabalhos, como: [10], [49], [50] e [6].

2.3 Decomposição de Séries Temporais Financeiras

O principal objetivo desta seção é introduzir os processos e transformações que serão aplicados às séries temporais para então analisar a possibilidade de incluí-las no modelo. Isso se torna necessário para entender melhor o processo de geração de tal série, ou seja, modelar o seu comportamento [21]. Com a série modelada, é possível realizar previsões cada vez mais acuradas sobre o comportamento futuro das séries temporais.

De forma a reduzir o efeito da não-estacionariedade e dos ruídos, fazendo com que as RNAs aprendam os padrões realmente relevantes [30], os seguintes processos

foram aplicados a todas as séries temporais consideradas neste estudo:

- Transformação em séries de retorno, pelo log da primeira diferença;
- Normalização dos dados;
- Extração de tendência linear;
- Extração de ciclos senoidais por autocorrelação e FFT;
- Atenuação de intrusos (clamping);
- Obtenção de estatísticas sobre os dados.

Por fim, são obtidas os valores que compõem as denominadas séries residuais. Tais séries apresentam o comportamento recomendado para inserção nas RNAs e podem ser facilmente reconstituídas para as séries originais, dada a aplicação dos operadores inversos.

2.3.1 Transformação em série de retorno

Como afirma Morettin [43], é comum encontrar diferentes níveis de média móvel em séries de preços de ativos financeiros, no decorrer destas. Uma das melhores saídas para evitar possíveis efeitos prejudiciais dessa tendência, ao inserir as séries nas RNAs, é trabalhar com séries de retornos, que refletem as variações diárias dos indicadores, em detrimento das séries de cotações.

Para efetuar a transformação, aplica-se a primeira diferença logarítmica na série, dada pela equação 2.17, onde r_t é o valor da série de retorno no tempo t e x_t é o valor da série original no tempo t . Com esta transformação, a série tem seus fatores de tendência reduzidos, ou seja, torna-se cada vez mais estacionária.

$$r_t = \log\left(\frac{x_t}{x_{t-1}}\right) \quad (2.17)$$

2.3.2 Normalização dos dados

Ao utilizar funções não-lineares na camadas da RNA, os valores que passam por seus neurônios geralmente precisam ser normalizados para corresponder ao intervalo de saídas destas funções [24].

Neste processo cria-se uma escala, estabelecendo um novo intervalo válido dentro do qual todos os dados são colocados. Os intervalos mais comumente utilizados na literatura são $[-1,1]$ ou $[0,1]$ e as funções mais utilizadas na normalização são:

- Transformação linear para o intervalo $[0,1]$, representada pela equação 2.18.

$$X_n = \frac{X_0 - X_{min}}{X_{max} - X_{min}} \quad (2.18)$$

- Transformação linear para o intervalo $[a,b]$, representada pela equação 2.19.

$$X_n = (b - a) \frac{X_0 - X_{min}}{X_{max} - X_{min}} + a \quad (2.19)$$

- Normalização estatística, representada pela equação 2.20.

$$X_n = \frac{X_0 - \bar{X}}{\sigma} \quad (2.20)$$

onde X_n é o valor normalizado, X_0 é o valor original, X_{max} é o valor máximo, X_{min} é o valor mínimo, \bar{X} é o valor médio e σ é o desvio padrão dos dados.

Mais tarde, na camada de saída, o resultado é denormalizado à sua forma original para a comparação com o valor desejado pela rede.

2.3.3 Extração de tendência

Tendência em uma série temporal é uma mudança sistemática e contínua em qualquer parâmetro de uma dada amostra, excluindo-se mudanças periódicas ou quase periódicas. Embora existam inúmeros tipos de tendência, esse trabalho se aterá apenas ao caso mais comum, chamada tendência linear.

O processo de extração de tendência linear é efetuado obtendo-se os parâmetros a e b da reta de regressão dos dados da série, na forma da equação 2.21 abaixo, e depois extraindo a série de tendência da série original.

$$\begin{aligned} tend(t) &= a.t + b \\ S(t) &= s(t) - tend(t) \end{aligned} \quad (2.21)$$

onde $s(t)$ é a série antiga e $S(t)$ é a nova série.

2.3.4 Extração de sazonalidades e ciclos senoidais

Sazonalidade pode ser definida como padrões de variação que ocorrem em intervalos regulares de tempo ou períodos. Um exemplo é a série do índice pluviométrico de uma determinada cidade no mundo, onde se sabe que, em média, em determinados meses há um aumento no volume de chuvas e em outros, ocorre estiagem. Entende-se por ciclo senoidal de uma série temporal o processo estacionário que exhibe um

movimento transitório em torno da tendência. É possível modelar a componente cíclica presente numa série utilizando o comportamento amortecido de uma senóide [43].

A literatura mostra que dois clássicos tratamentos de séries temporais podem ser aplicados às séries estudadas, com o objetivo de buscar sazonalidades e ciclos senoidais, fenômenos frequentemente presentes em séries desse tipo. A análise de Fourier usualmente é realizada através da transformada de Fourier, pelo algoritmo Fast Fourier Transform (FFT), que tem como saída os parâmetros a e b da equação de extração de ciclos senoidais, definida pela equação 2.22 abaixo.

$$\begin{aligned} cs(t) &= a * \cos(2\pi ft) + b * \sin(2\pi ft) \\ S(t) &= s(t) - cs(t) \end{aligned} \tag{2.22}$$

Por sua vez, a autocorrelação de uma série temporal é definida pela correlação entre a série $y(t)$ atual e a mesma série atrasada de k unidades de tempo, $y(t - k)$. Caso disponham-se de N valores de uma série estacionária no tempo, podem-se montar $N - k$ pares $[y(t), y(t + k)]$, $t = 1, \dots, N - k$.

O objetivo é verificar se pode ser caracterizada alguma autocorrelação significativa em atrasos $P, 2P, 3P, \dots$, incluindo a evidência no espectrograma com raias em $1/P, 2/P, 3/P, \dots$. Se houver frequências altas no espectrograma e a autocorrelação apresentar um aspecto de senóide, há evidências para a remoção de ciclos senoidais.

2.3.5 Atenuação de intrusos (*clamping*)

Assim como no trabalho de Constant [10], foi utilizado aqui um tratamento que consiste em minimizar o efeito dos pontos extremos das séries, uma vez que a rede prioriza a redução dos maiores erros. Tais valores são caracterizados por se apresentarem acima ou abaixo de um limiar calculado a cada ponto da série, considerando valores atrasados.

Dado um ponto em D na série, obter os $D - 1$ a $D - 10$ pontos (total de 10 valores) para calcular sua média μ e desvio padrão σ . Caso o valor em D seja maior que $(\mu + n * \sigma)$ ou menor que $(\mu - n * \sigma)$, a série é atenuada para tal valor limite. Foi utilizado o valor de $n = 2$ desvios-padrões para os experimentos.

2.3.6 Estatísticas

As estatísticas extraídas dos dados de análise são descritas nesta seção.

- Média: valor médio de uma população usualmente utilizado como valor da tendência central de um conjunto de dados, para valores reais de variáveis

aleatórias a média representa o valor esperado destas variáveis;

- Mediana: também conhecido como percentil de 50%, é o valor que separa uma população onde metade das amostras é superior e metade é inferior ao valor estimado;
- Variância: em estatística é uma medida da dispersão de uma variável aleatória, indicando quão distante as variáveis estão do valor esperado de uma distribuição;
- Desvio Padrão: é a medida mais utilizada para indicar como os dados estão dispersos em um conjunto e definido como a raiz quadrada da variância;
- Obliquidade (*skewness*): é uma medida da assimetria da distribuição de valores reais de uma variável aleatória, mais precisamente seu valor será positivo quando a cauda superior for mais longa e negativo no caso contrário;
- Curtose (kurtosis): é uma medida de dispersão que caracteriza o "achatamento" de uma distribuição. Um alto coeficiente indica que boa parte da variância é devido a desvios pouco frequentes, ou seja, um pico suave e caudas mais "gordas".

2.4 Estado da Arte

Há uma grande diversidade de trabalhos na literatura que estudam o problema de predição no mercado financeiro. Os problemas são abordados por diferentes técnicas, que variam de métodos estatísticos a avançados modelos utilizando *soft computing*, inclusive com modelos híbridos, misturando tais técnicas [6].

Quando observa-se a quantidade de trabalhos que investigam o mercado brasileiro, nota-se um número reduzido de pesquisa, vide o trabalho de pesquisa de Atsalakis & Valavanis [6], onde apenas um dos trabalhos analisados realiza predições sobre o Ibovespa.

Zhang [24] apresenta um dos mais completos *surveys* na área de predição de séries temporais, referência até hoje em diversos aspectos de arquitetura, topologia e metodologia de implementação de RNAs com esta finalidade, mesmo sendo um trabalho de 1998. Nele, Zhang apresenta excelente estudo comparativo sobre as conclusões que os trabalhos analisados descrevem. Ele afirma ainda que, apesar de uma quantidade considerável de pesquisa ter sido realizada nesta área, seus resultados não são conclusivos sobre se e quando RNAs são melhores do que os métodos clássicos para predição.

O trabalho de Atsalakis e Valavanis [6] também é um grande *survey* de aplicações de RNA na previsão de séries temporais financeiras, em uma época mais recente (2009). Nele são reportados mais de cem artigos de diversos autores, e seus resultados são apresentados em cinco tabelas, contendo filtros por: mercado de ações aplicado, variáveis de entrada, metodologia e parâmetros utilizados, estudos comparativos com diferentes arquiteturas e, por fim, as medidas de performance extraídas desses modelos. Também se enquadram nesse aspecto de bons *surveys* os trabalhos de Palywal [50] e Krollner [33].

Mais recentemente, podem ser encontrados trabalhos de diversos autores que buscam realizar as predições e classificações das mais variadas formas. Martinez [34] realiza a predição de *high* e *low* e utiliza os valores estimados em um *Trading System*, com diversas restrições reais do mercado. Em uma linha similar, Gomide [26] realiza um trabalho de predição sobre o *spread* entre ações do mesmo setor, também aplicando os resultados em um sistema de negociação automatizado, em formato de negociação em par (*pair trading*). Outro trabalho que aplica os resultados em *pair trading* é a dissertação de Mestrado de Madhavaram [37], que utiliza SVM em uma estratégia sobre ações do setor financeiro e o fundo de índice (ETF) que as representa, chamado XLF.

Como exemplo recente de estudo comparativo, pode-se citar o trabalho de Agrawal & Murarka [46], que realiza a predição do fechamento em $D + 1$ para diferentes empresas indianas pertencentes a distintos setores da economia, como: TI, automobilismo e bancos. São realizadas comparações entre diferentes tipos de RNAs, como *Backpropagation Neural Network* (BPNN), *Radial Basis Function Neural Network* (RBFNN), *Generalized Regression Neural Network* (GRNN) e *Layer Recurrent Neural Network* (LRNN). Já o trabalho de Mantri [40], realiza a comparação de seus resultados de predição por RNAs MLP e SVM.

Há trabalhos que utilizam os resultados das RNAs como parte de um sistema maior, que envolve outros indicadores para guiar a disparada de ordens de compra e venda. Em Barroso [11], por exemplo, é realizada uma busca por gatilhos de compra e venda de ativos financeiros, baseada em *candlesticks* de Análise Técnica, aliado a uma estratégia com RNAs e simulando as operações nos momentos indicados pelo algoritmo. Em [56], RNAs são utilizadas na predição de valores futuros de ativos no mercado de câmbio Forex (FX). Os resultados são utilizados em uma estratégia automatizada de forma a evitar o envio de ordens que, dada a estratégia já existente, não sejam consistentes com o resultado da rede.

Outros tipos de métodos são também hibridizados com as RNAs. É o caso, por exemplo, dos trabalhos de Anderson [20] e ORTEGA [48], que se utilizam das teorias de Wavelets para processar as séries temporais, antes de utilizar como entrada na RNA.

Os dados de entrada destes trabalhos podem vir de diversas fontes. Podem ser índices financeiros mundialmente conhecidos, como: S&P, Nasdaq, DAX, não limitados a estes. Outros exemplos são os indicadores oriundos da análise técnica [33], onde são analisados indicadores de tendências, reversão de tendências, resistências, suportes e movimentos baseados em volume, não limitados a estes, mas sempre baseados na série histórica de preços do ativo. Ainda, podem-se incluir na predição valores referentes à análise fundamentalista [20], como dados macroeconômicos - tais como índices de inflação, desemprego, PIB etc. - e dados de balanços trimestrais ou anuais das empresas.

É possível notar a superioridade de trabalhos que utilizam modelos MLP com algoritmo de treinamento *backpropagation*, características que foram também adotadas neste trabalho. O diferencial são as variáveis de saída adotadas, onde, ao invés de se utilizar a clássica predição do preço de fechamento ou a decisão de compra ou venda do ativo, neste trabalho o foco é também a predição de variáveis que auxiliem estratégias intradiárias de investimento, como *high*, *low* e *order*.

Apesar da enorme diversidade de modelos e métodos, nenhum destes trabalhos reporta extraordinários retornos. Uma possibilidade para tal fenômeno é teorizada por Granger [27], que sugere um possível "viés de informação", onde se um método de previsão que obtém lucros constantes é encontrado, o pesquisador preferirá lucrar com ele a publicar o estudo.

Capítulo 3

Mercado Financeiro

3.1 Estrutura do mercado financeiro

O termo "mercado" pode ser compreendido como o processo pelo qual pessoas interessadas em vender algum produto ou serviço se encontram com pessoas interessadas em comprar esse mesmo produto ou serviço. Logo, o mercado financeiro pode ser definido como o lugar onde os ativos financeiros são negociados.

Com o processo de globalização dos mercados financeiros internacionais, os países em desenvolvimento procuram abrir suas economias para poder receber investimentos externos. Os principais objetivos são permitir o desenvolvimento das empresas e se constituir como uma importante opção de investimento para pessoas e instituições.

Há, basicamente, dois lugares onde estes ativos são negociados, que são os mercados regulamentados pelas bolsas de valores e os mercados de balcão. As bolsas de valores são os locais que oferecem as condições e os sistemas necessários para a realização de negociação de compra e venda de títulos e valores mobiliários, e de outros ativos, de forma transparente [8].

No Brasil, a principal bolsa de valores é a BM&FBovespa e nela são negociados ativos de diferentes classes, como: ações, opções de ação, contratos de índice, contratos de futuros de commodities, opções sobre contratos futuros, operações a termo, entre diversos outros ativos. Também variadas são as expectativas dos investidores, que podem ser, por exemplo: acumulação de capital, especulação, alavancagem de patrimônio, proteção contra variações de câmbio, dentre outras.

O Sistema Financeiro Nacional (SFN) é composto de instituições responsáveis pela captação de recursos financeiros, pela distribuição e circulação de valores e pela regulação desse processo [8]. O Conselho Monetário Nacional (CMN), seu organismo maior, presidido pelo ministro da Fazenda, é quem define as diretrizes de atuação do sistema. Diretamente ligados a ele estão o Banco Central do Brasil, que atua como

seu órgão executivo, e a Comissão de Valores Mobiliários (CVM), que responde pela regulamentação e fomento do mercado de valores mobiliários (de bolsa e de balcão).

3.1.1 Histórico da BM&FBOVESPA

A BM&FBOVESPA foi criada em 8 de maio de 2008, com a integração entre a Bolsa de Mercadorias & Futuros (BM&F, bolsa de derivativos), criada em 1917, e a Bolsa de Valores de São Paulo (BOVESPA, bolsa de ações), fundada em 1890. Juntas, as companhias presentes formam uma das maiores bolsas do mundo em valor de mercado - a décima maior do mundo em capitalização de mercado (US\$ 1,5 trilhão), de acordo com a Forbes [1] -, a segunda das Américas e a líder na América Latina. Atualmente, a BM&FBOVESPA é o maior centro de negociação com ações da América Latina, destaque que culminou com um acordo histórico para a integração de todas as bolsas brasileiras em torno de um único mercado de valores.

Os mercados disponíveis na BM&FBOVESPA são [2]:

- Mercado à vista: No qual a liquidação física (entrega de títulos vendidos) se processa no 2º dia útil após a realização do negócio em bolsa e a liquidação financeira (pagamento e recebimento do valor da operação) se dá no 3º dia útil posterior à negociação, e somente mediante a efetiva liquidação física.
- Mercado a Termo: Operações com prazos de liquidação diferidos, em geral, de trinta, sessenta ou noventa dias. O contrato a termo pode, ainda, ser liquidado antes de seu vencimento.
- Mercado de Opções: A opção proporciona ao investidor que a adquire, mediante o pagamento de um prêmio em dinheiro, o direito de comprar ou vender um lote de ações a outro investidor, com preço e prazo de exercício preestabelecido contratualmente.

As principais empresas com capital aberto na BM&FBOVESPA são as que compõem seu principal índice, o Ibovespa, objeto de estudo dessa dissertação. Estudar o índice significa compreender melhor sobre o mercado financeiro que se deseja investir, no nosso caso, o brasileiro.

3.1.2 Índice Ibovespa

O Ibovespa é o resultado de uma carteira teórica de ativos, elaborada de acordo com os critérios estabelecidos em sua metodologia, que está descrita no Manual de Definições e Procedimentos dos Índices da BM&FBOVESPA [2]. O Ibovespa foi criado em 1968, com a finalidade básica de servir como indicador médio do comportamento do mercado.

O índice representa o valor atual, em moeda corrente, de uma carteira teórica de ações constituída em 02/01/1968, a partir de uma aplicação hipotética. Supõe-se não ter sido efetuado nenhum investimento adicional desde então, considerando-se somente os ajustes efetuados em decorrência da distribuição de proventos pelas empresas emissoras (tais como reinversão de dividendos recebidos e do valor apurado com a venda de direitos de subscrição, e manutenção em carteira das ações recebidas em bonificação) [10]. Dessa forma, o índice reflete não apenas as variações dos preços das ações, mas também o impacto da distribuição dos proventos, sendo considerado um indicador que avalia o retorno total de suas ações componentes. É importante perceber que não estão incluídos no índice as "BDRs¹ e ativos de companhias em recuperação judicial ou extrajudicial, regime especial de administração temporária, intervenção ou que sejam negociados em qualquer outra situação especial de listagem" [2].

Atualmente, o índice é composto por 73 ativos entre ações e Units², listados na tabela 3.1. As informações contém o código do ativo, o nome da empresa, o tipo da ação (preferencial ou ordinária) e porcentagem de participação no índice, divididas em duas partes, lado a lado.

3.2 Séries temporais financeiras

Um grupo de séries temporais que tem sido alvo de intenso estudo por parte de economistas, estatísticos, matemáticos, engenheiros e outros profissionais são as séries financeiras. Este grupo de séries é composto principalmente pelas séries de cotações em mercados de ações, mas engloba também séries de taxas de câmbio e outras fontes de séries.

A previsão sobre séries temporais financeiras é bastante difícil porque os sistemas que geram os valores das séries são extraordinariamente complexos e a informação disponível para realizar as previsões é limitada. O processamento e modelagem de séries temporais financeiras são tarefas comumente realizadas por pesquisadores e empresas, com os mais variados propósitos: predição de preços, estimação da volatilidade de ativos, análise de risco de mercado, otimização de portfólios de investimento, dentre várias outras.

¹BDRs (*Brazilian Depositary Receipts*) são certificados de depósito de valores mobiliários emitidos no Brasil que representam valores mobiliários de emissão de companhias abertas com sede no exterior.

²Units são ativos compostos por mais de uma classe de valores mobiliários, como uma ação ordinária e um bônus de subscrição, por exemplo, negociados em conjunto e compradas/vendidas como uma unidade.

Código	Ação	Tipo	Part. (%)	Código	Ação	Tipo	Part. (%)
ABEV3	AMBEV S/A	ON	4,214	GGBR4	GERDAU	PN ED N1	1,775
AEDU3	ANHANGUERA	ON NM	0,794	GOAU4	GERDAU MET	PN ED N1	0,524
ALLL3	ALL AMER LAT	ON NM	0,627	GOLL4	GOL	PN N2	0,382
BBAS3	BRASIL	ON NM	2,492	HGTX3	CIA HERING	ON NM	0,667
BBDC3	BRADESCO	ON EJ N1	1,491	HYPE3	HYPERMARCAS	ON NM	0,868
BBDC4	BRADESCO	PN EJ N1	5,754	ITSA4	ITAUSA	PN ED N1	2,903
BBSE3	BBSEGURIDADE	ON NM	1,720	ITUB4	ITAUUNIBANCO	PN ED N1	7,208
BISA3	BROOKFIELD	ON NM	0,337	JBSS3	JBS	ON NM	1,227
BRAP4	BRADESPAR	PN N1	0,619	KLBN11	KLABIN S/A	UNT N2	0,352
BRFS3	BRF SA	ON NM	2,582	KLBN4	KLABIN S/A	PN N2	0,352
BRKM5	BRASKEM	PNA N1	0,549	KROT3	KROTON	ON NM	1,206
BRML3	BR MALLS PAR	ON NM	1,337	LAME4	LOJAS AMERIC	PN	0,672
BRPR3	BR PROPERT	ON NM	0,772	LIGT3	LIGHT S/A	ON NM	0,250
BVMF3	BMFBOVESPA	ON NM	2,633	LLXL3	LLX LOG	ON NM	0,189
CCRO3	CCR SA	ON NM	1,726	LREN3	LOJAS RENNER	ON NM	0,984
CESP6	CESP	PNB N1	0,455	MFRG3	MARFRIG	ON NM	0,397
CIEL3	CIELO	ON NM	2,308	MRVE3	MRV	ON NM	0,743
CMIG4	CEMIG	PN N1	1,311	NATU3	NATURA	ON NM	0,916
CPFE3	CPFL ENERGIA	ON NM	0,541	OIBR4	OI	PN N1	0,789
CPLE6	COPEL	PNB N1	0,334	PCAR4	P.ACUCAR-CBD	PN N1	1,451
CRUZ3	SOUZA CRUZ	ON	0,838	PDGR3	PDG REALT	ON NM	0,891
CSAN3	COSAN	ON NM	0,660	PETR3	PETROBRAS	ON	3,653
CSNA3	SID NACIONAL	ON ED	0,973	PETR4	PETROBRAS	PN	7,338
CTIP3	CETIP	ON NM	0,893	QUAL3	QUALICORP	ON NM	0,520
CYRE3	CYRELA REALT	ON NM	0,707	RENT3	LOCALIZA	ON NM	0,631
DASA3	DASA	ON NM	0,202	RSID3	ROSSI RESID	ON NM	0,345
DTEX3	DURATEX	ON NM	0,402	SANB11	SANTANDER BR	UNT N2	1,166
ECOR3	ECORODOVIAS	ON NM	0,410	SBSP3	SABESP	ON NM	0,835
ELET3	ELETROBRAS	ON N1	0,260	SUZB5	SUZANO PAPEL	PNA N1	0,754
ELET6	ELETROBRAS	PNB N1	0,380	TBLE3	TRACTEBEL	ON NM	0,640
ELPL4	ELETROPAULO	PN N2	0,292	TIMP3	TIM PART S/A	ON NM	1,136
EMBR3	EMBRAER	ON NM	1,501	UGPA3	ULTRAPAR	ON ED NM	1,456
ENBR3	ENERGIAS BR	ON NM	0,351	USIM5	USIMINAS	PNA N1	0,930
ESTC3	ESTACIO PART	ON NM	0,867	VALE3	VALE	ON N1	3,809
EVEN3	EVEN	ON NM	0,335	VALE5	VALE	PNA N1	7,680
FIBR3	FIBRIA	ON NM	0,776	VIVT4	TELEF BRASIL	PN	1,296
GFS3	GAFISA	ON NM	0,622	Quantidade Teórica Total			100,000

Tabela 3.1: Composição atual do índice Ibovespa

3.2.1 Séries Temporais consideradas neste trabalho

No total, foram obtidas e estudadas 43 possíveis séries temporais de entrada. A ideia de considerar cada uma destas séries é originária de outros trabalhos na literatura, como por exemplo, [6] [50] [33] [20], e da opinião de especialistas e analistas econômicos. A seguir, é exibida uma listagem resumida com todas as 43 séries, subdivididas em algumas categorias de interesse para este trabalho.

- Séries do próprio Ibovespa (19):

Open, High, Low, Close, Order;

Feriados nos EUA, Véspera de Feriados no Brasil, Variação Diária, Variação Máxima, Volume;

Análise Técnica: IFR, MM (10), MM(132), MACD, OBV, *Bollinger Bands*, IVR, Momento, Volatilidade Histórica;

- Americas (6):

Argentina, Canadá, Chile, México, EUA (*Dow Jones Industrial Average* e S&P500);

- Ásia (6):

Hong Kong, China, Índia, Japão, Coreia do Sul, Rússia;

- Europa (4):

Inglaterra, França, Alemanha, Itália;

- Outras séries importantes (8)

CDI, Dólar PTAX, Euro, FGV100, IDIV, IMOB, SMLL, IND\$ (fut. de Ibovespa).

A seguir, é exibida uma breve explicação sobre o significado de cada série temporal financeira considerada neste trabalho, com a exibição dos gráficos correspondentes a estas séries. As variáveis de análise técnica serão explicadas e exibidas com maiores detalhes na seção 3.3.2.

Séries do próprio Ibovespa

1. Séries diárias de preços de fechamento, de abertura, valor máximo e valor mínimo do índice, representado por *close*, *open*, *high* e *low*, exibidos na figura 3.1.

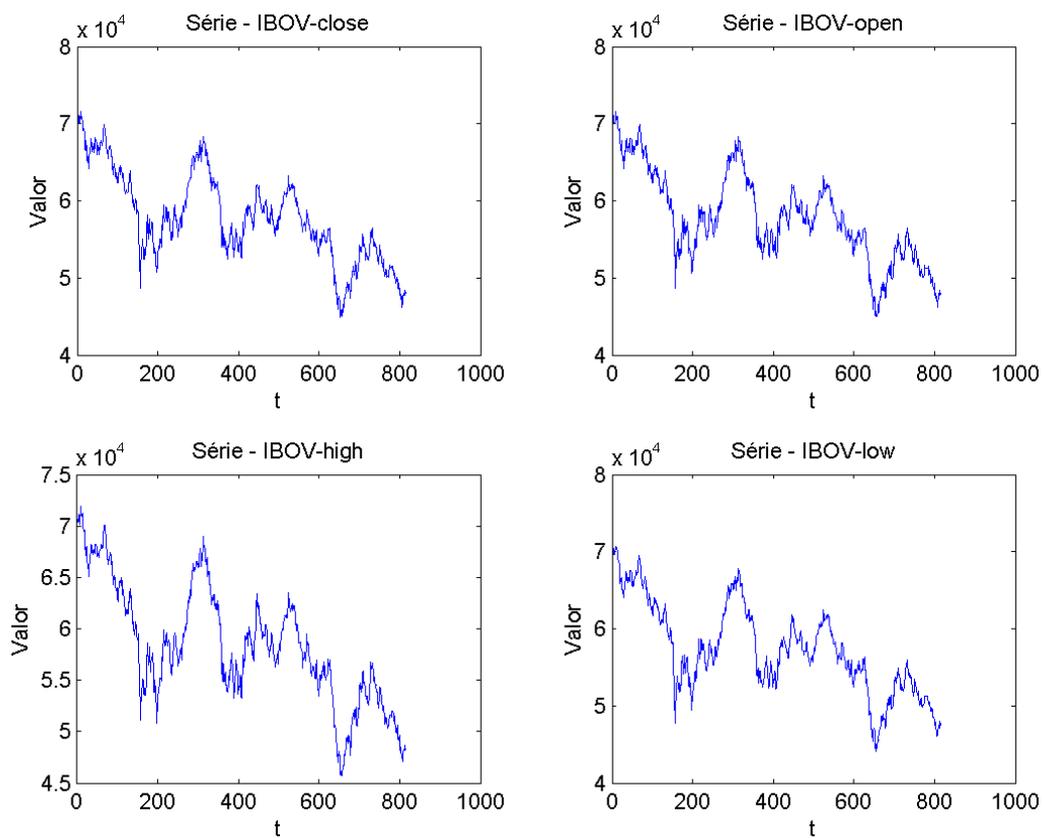


Figura 3.1: Gráficos com as séries diárias de *close*, *open*, *high* e *low* do Ibovespa

2. Série diária do *order*, a ordem em que *high* e *low* ocorrem no período intradiário, na figura 3.2.

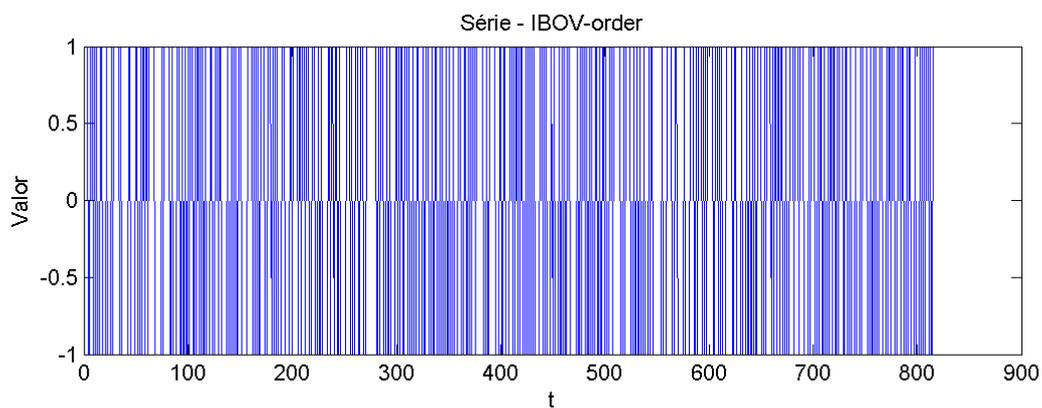


Figura 3.2: Gráfico da série diária do *order*

3. Série diária do volume de negociação do Ibovespa, representada pela figura 3.3.

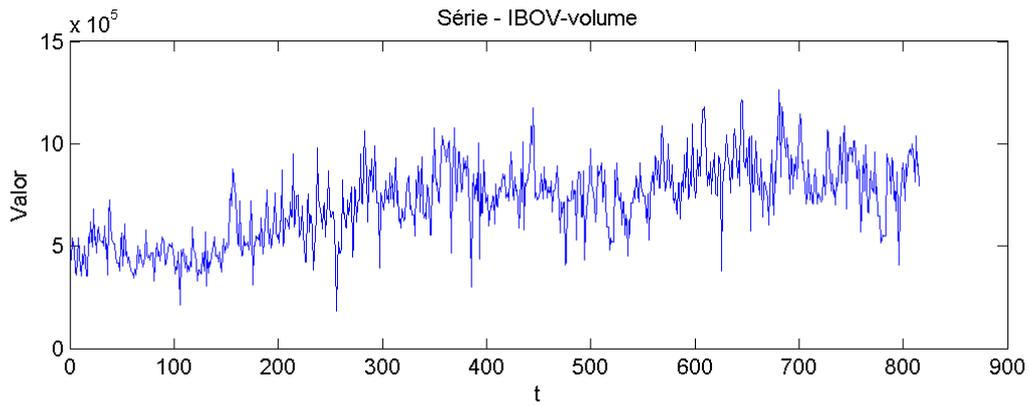


Figura 3.3: Gráfico com a série de Volume do Ibovespa

4. Feriados nos EUA, Vésperas de Feriado no Brasil: busca por algum padrão específico nesses dias, quando a variação de preços é atípica dado o menor volume de negociação. Os gráficos com as representações destas séries são exibidos na figura 3.4.

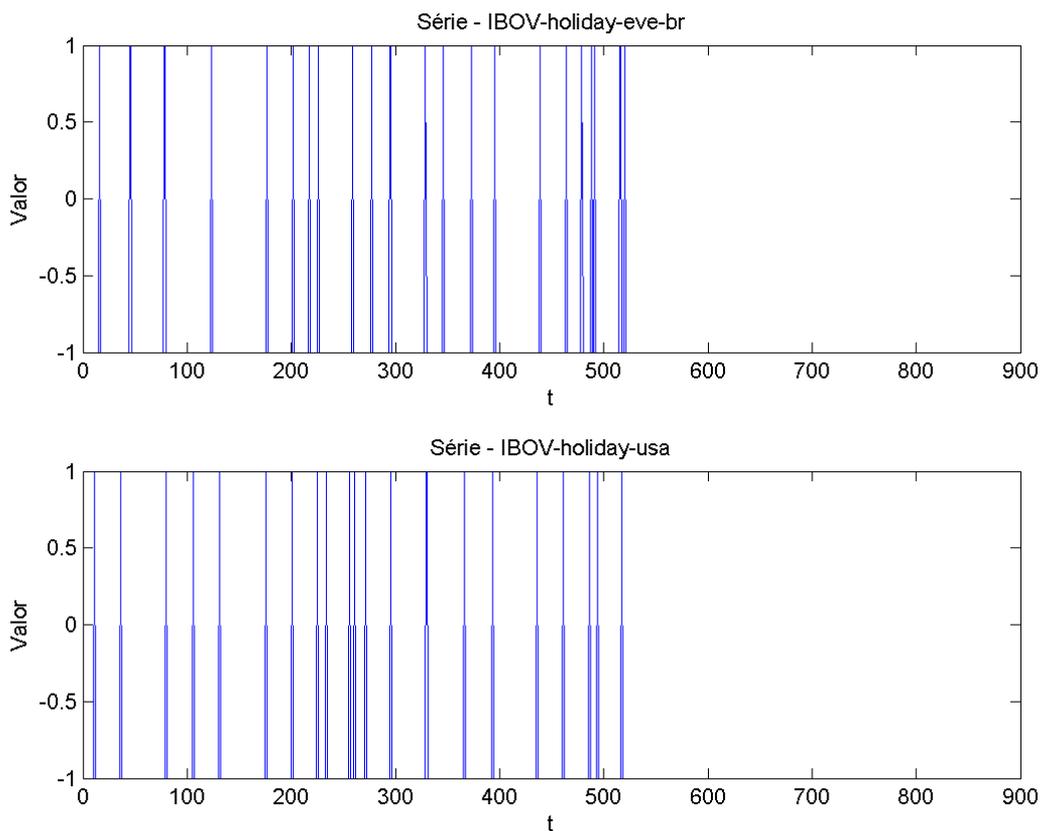


Figura 3.4: Gráficos das séries diárias de Feriado nos EUA e Vésperas de Feriado no Brasil

5. Além destas, foram extraídas mais duas importantes séries, já em forma de retorno, de forma a incrementar a investigação sobre as entradas significativas

do modelo. Ambas as séries foram também incluídas no método proposto em Constant [10]. As duas novas séries são: Variação Diária (VarDiaria), que representa a variação diária dos preços, entre fechamento e abertura; e Variação Máxima (VarMaxima), que representa a máxima variação intradiária do Ibovespa. Suas fórmulas são dadas pelas equações 3.1 e 3.2 e seus gráficos estão na figura 3.5, ambos exibidos a seguir.

$$VarDiaria(t) = \log\left(\frac{close(t)}{open(t)}\right) \quad (3.1)$$

onde $close(t)$ representa o valor de fechamento do Ibovespa em t e $open(t)$ representa o preço de abertura em t .

$$VarMaxima(t) = \log\left(\frac{high(t)}{low(t)}\right) \quad (3.2)$$

onde $high(t)$ e $low(t)$ representam, respectivamente, os preços máximo e mínimo do Ibovespa em t .

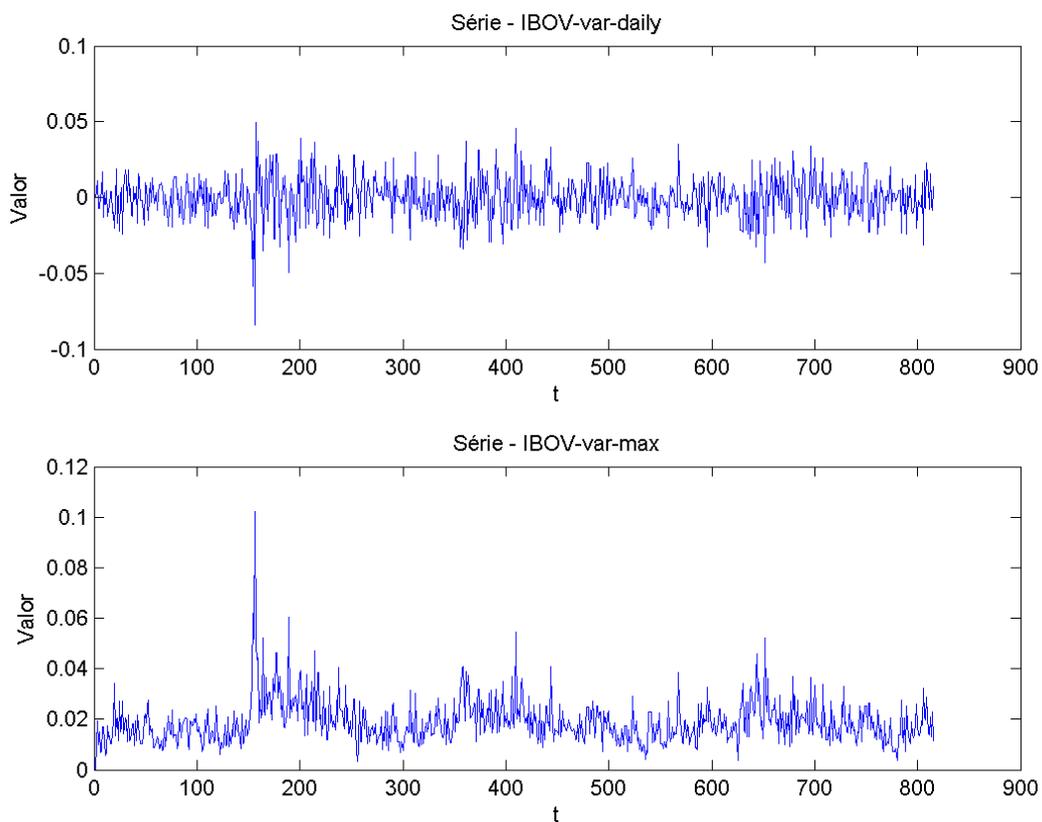


Figura 3.5: Gráficos das séries de Variação diária e variação máxima do Ibovespa

Índices mundiais - Américas

De forma a capturar a influências dos mercados externos nas variações internas do Ibovespa, foram obtidas as séries de valores para diversos índices de bolsas mundiais. A listagem completa é exibida a seguir, categorizada pela região de origem.

1. Argentina Merval Index (*MER*cado de *VAL*ores): é a marcação a mercado de uma cesta de papéis selecionados de acordo com sua liquidez na Buenos Aires Stock Exchange.
2. O Canada S&P/TSX Index é um índice dos preços das ações das maiores empresas na Bolsa de Toronto (TSX) em termos de capitalização de mercado, correspondendo a cerca de 70% da capitalização de mercado de todas as empresas listadas no TSX.

Uma representação das séries Argentina Merval e Canadá S&P/TSX Index é exibida na figura 3.6 a seguir.

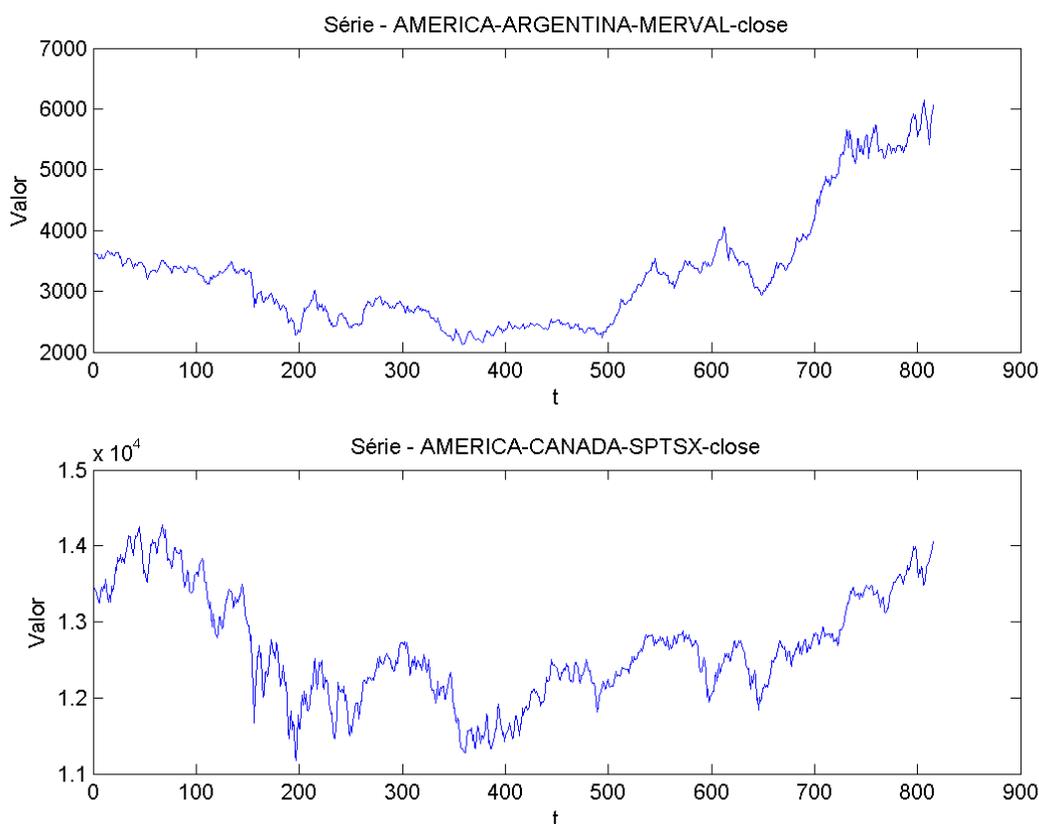


Figura 3.6: Gráficos das séries diárias dos índices da Argentina e do Canadá

3. O Índice de Preços Seletivos de Ações (IPSA Index), do Chile, é composto pelas 40 empresas com as maiores médias anuais de negociação listadas na bolsa de ações de Santiago.

4. O índice IPC, da bolsa de valores do México, é um índice ponderado pelas maiores empresas de capital aberto do país.

Uma representação contendo as séries Chile IPSA e México IPC é exibida na figura 3.7 a seguir.

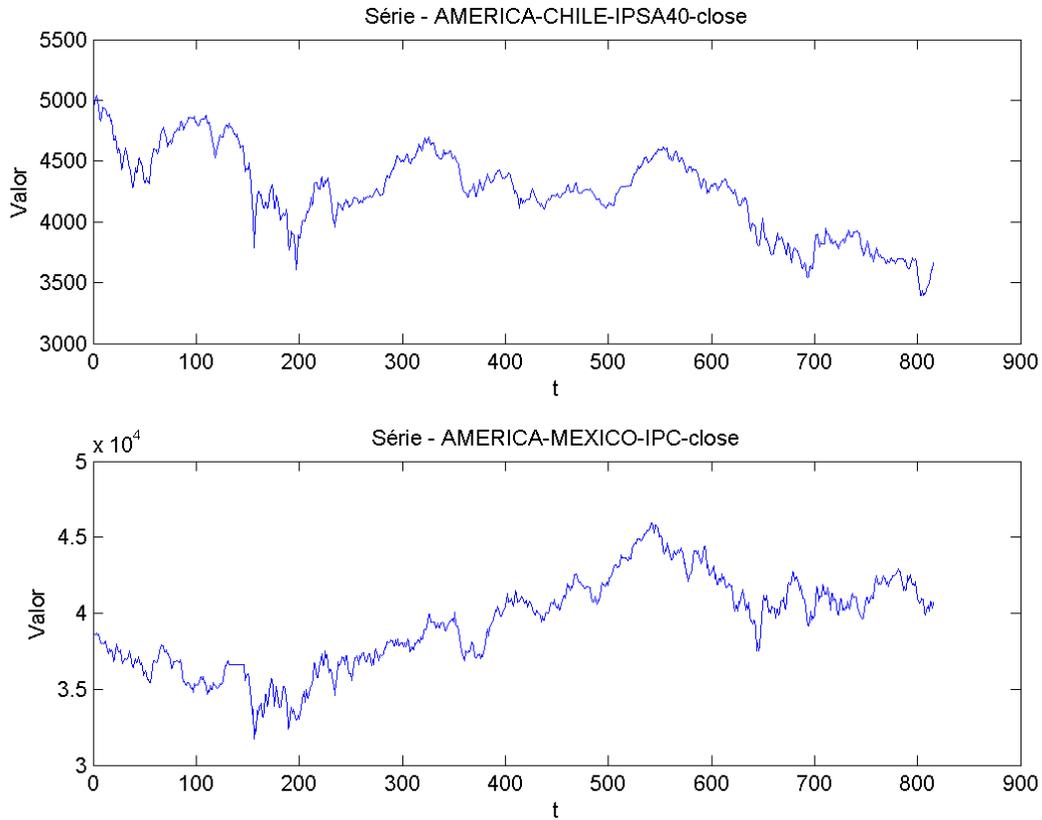


Figura 3.7: Gráficos das séries diárias dos índices de Chile e México

5. O índice Dow Jones Industrial Average (DJIA) é um índice criado em 1896 pelo editor do The Wall Street Journal e fundador do Dow Jones & Company, Charles Dow. É um dos principais indicadores dos movimentos do mercado americano, sendo baseado na cotação das ações das 30 maiores empresas do setor industrial americano.
6. O índice Standard and Poor's 500, comumente chamado de S&P 500, é um índice ponderado pelas 500 maiores empresas listadas nas bolsas americanas NYSE e Nasdaq.

A imagem com as séries dos EUA, contendo os índices DJIA e S&P 500, é exibida na figura 3.8 a seguir.

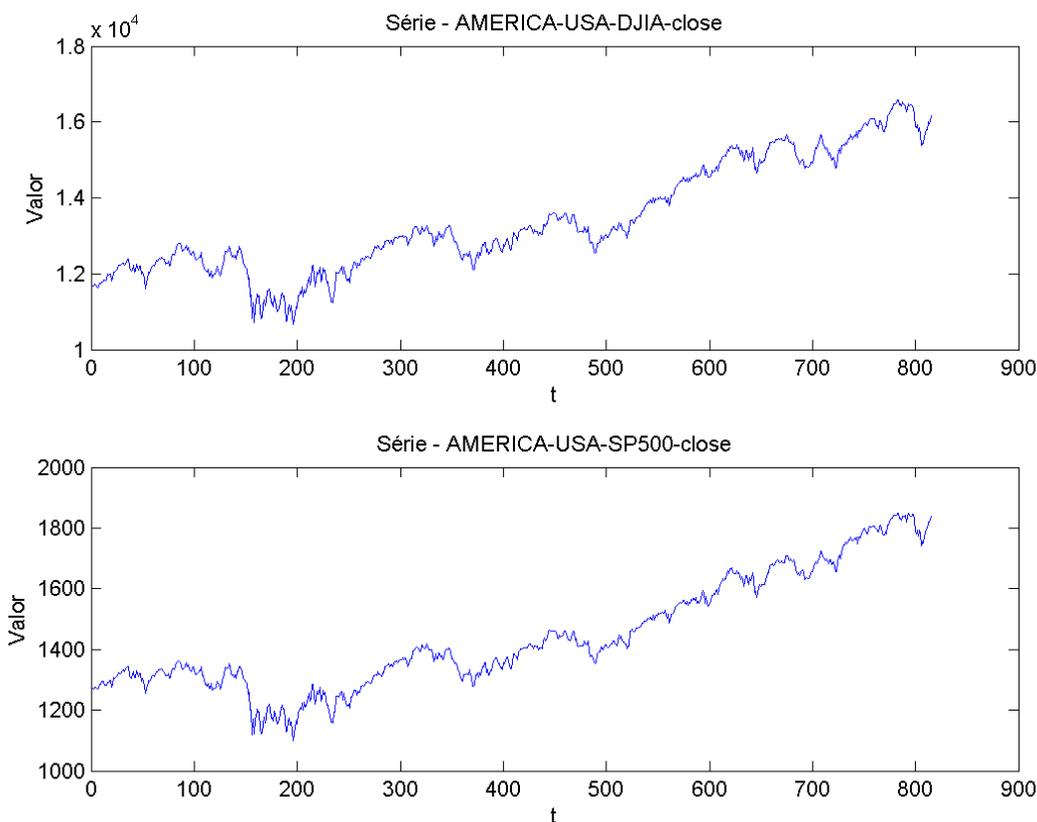


Figura 3.8: Gráficos das séries diárias dos índices americanos: DJIA e S&P 500

Índices mundiais - Ásia

1. O índice Shanghai Stock Exchange Composite (SSECO) é constituído por ações listadas na bolsa de Xangai, considerando a capitalização total de mercado de todas as ações daquele dia.
2. Hang Seng Index é o índice de ações da bolsa de Hong Kong, que considera 48 ações das maiores empresas listadas, correspondendo a aproximadamente 60% da capitalização do mercado.

A figura 3.9, a seguir, exhibe os valores diários para as séries de índices de Xangai e Hong Kong.

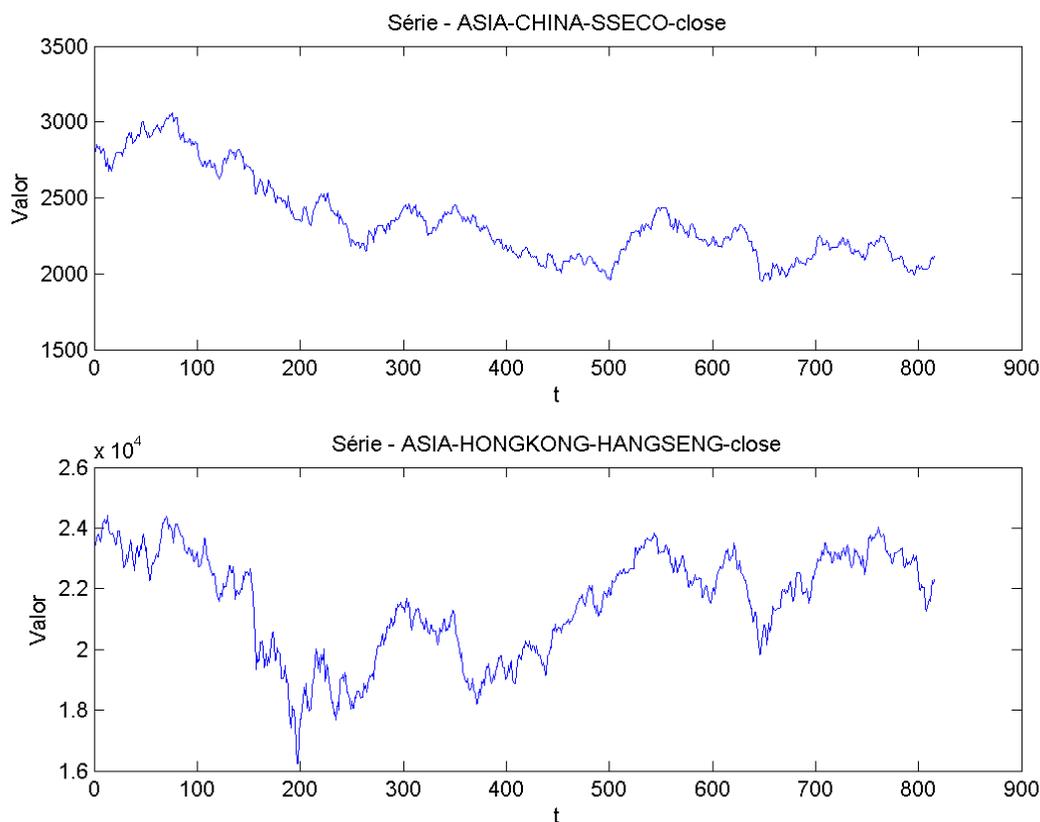


Figura 3.9: Gráficos das séries diárias dos índices da China e de Hong Kong

3. O S&P BSE SENSEX (S&P Bombay Stock Exchange Sensitive Index), também denominado BSE 30 ou simplesmente SENSEX, é o índice ponderado de ações da bolsa indiana de Mumbai. É formado pelos preços das 30 empresas mais estáveis e com melhor saúde financeira, listados em tal bolsa.
4. O índice Nikkei 225 é o mais conhecido dos índices asiáticos e representa o mercado acionário da Tokyo Stock Exchange, bolsa japonesa. Ele é concebido para refletir a saúde econômica do mercado japonês, desse modo, agrupando 225 empresas de quase todos os setores da economia.

A figura 3.10, por sua vez, apresenta a série de preços dos índices S&P BSE SENSEX e Nikkei 225.

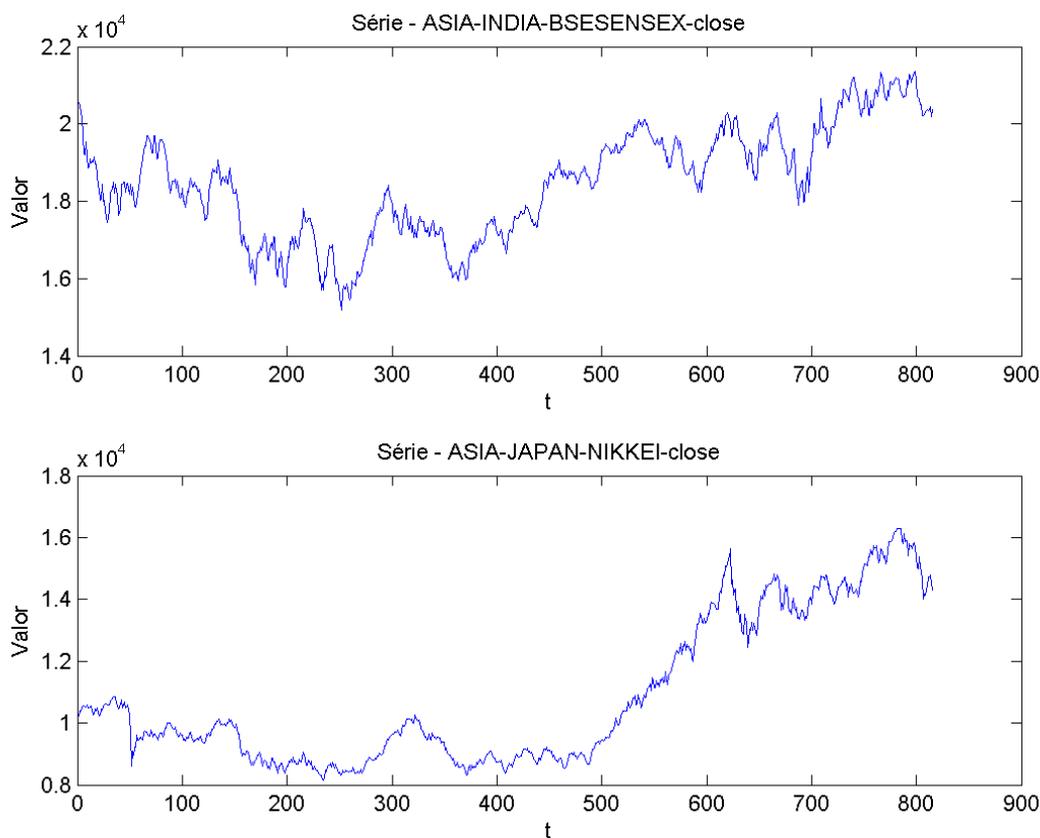


Figura 3.10: Gráficos das séries diárias dos índices da Índia e do Japão

5. Diferente de outros índices que limitam o número de ações, seja selecionando pela liquidez ou pela capitalização da empresa, o Korea Composite Stock Price Index (KOSPI) é composto por todas as ações negociadas na Korea Stock Exchange. Ele apenas as pondera de acordo com sua capitalização no mercado [49].
6. O RTS Index (RTSI) é um índice ponderado pela capitalização de 50 ações russas negociadas na Bolsa de Moscou, na Rússia. A lista de ações (aproximadamente 50) é revista a cada três meses pelo comitê de mercado russo. A figura 3.11 exhibe, portanto, os preços diários dos índices KOSPI e RTSI.

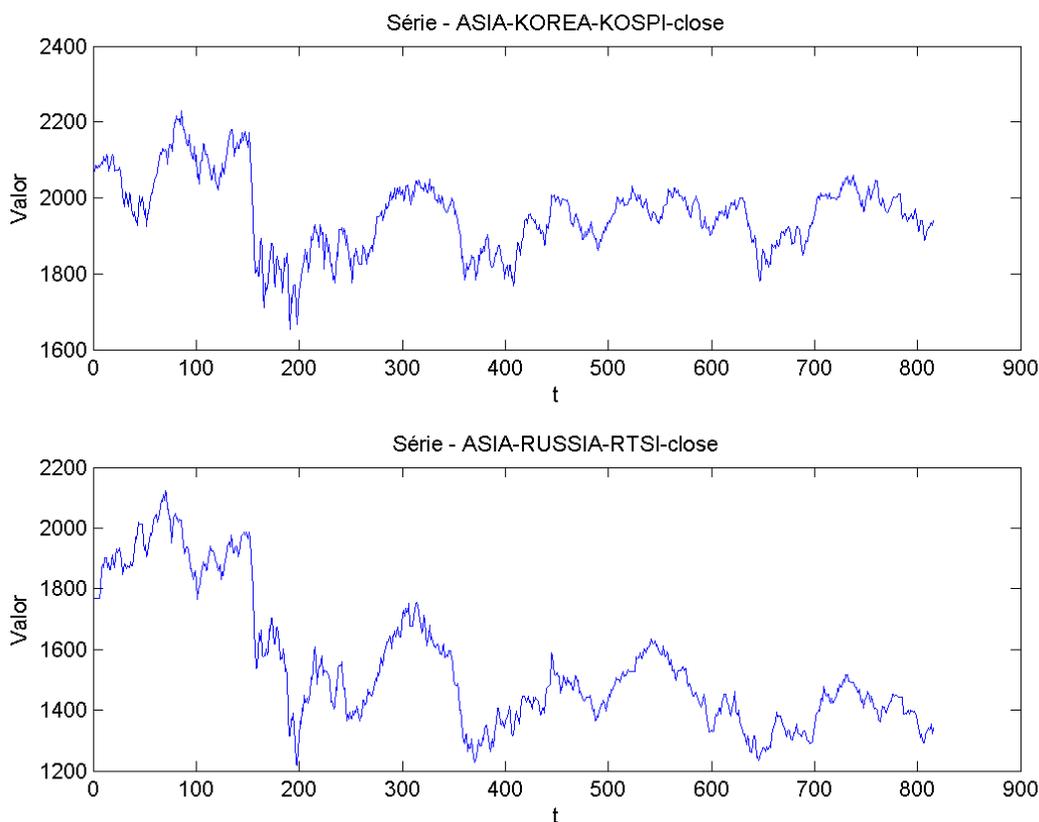


Figura 3.11: Gráficos das séries diárias dos índices da Coreia do Sul e da Rússia

Índices mundiais - Europa

1. O CAC 40 é o índice da Bolsa de Paris (Euronext Paris), constituído pelas 40 empresas com a maior capitalização bursátil, sendo a data-base de contabilização de seu valor estabelecida em 31 de dezembro de 1987.
2. O DAX 30 (*Deutscher Aktien-Index*) é o índice das 30 maiores empresas negociadas na Frankfurt Stock Exchange, na Alemanha, e um dos principais da Europa e do mundo. Esse índice também é conhecido por ser composto apenas por empresas blue chips, ou seja, com alto valor de mercado.

Uma representação com as séries do CAC 40 e o DAX 30 é exibida na figura 3.12 a seguir.

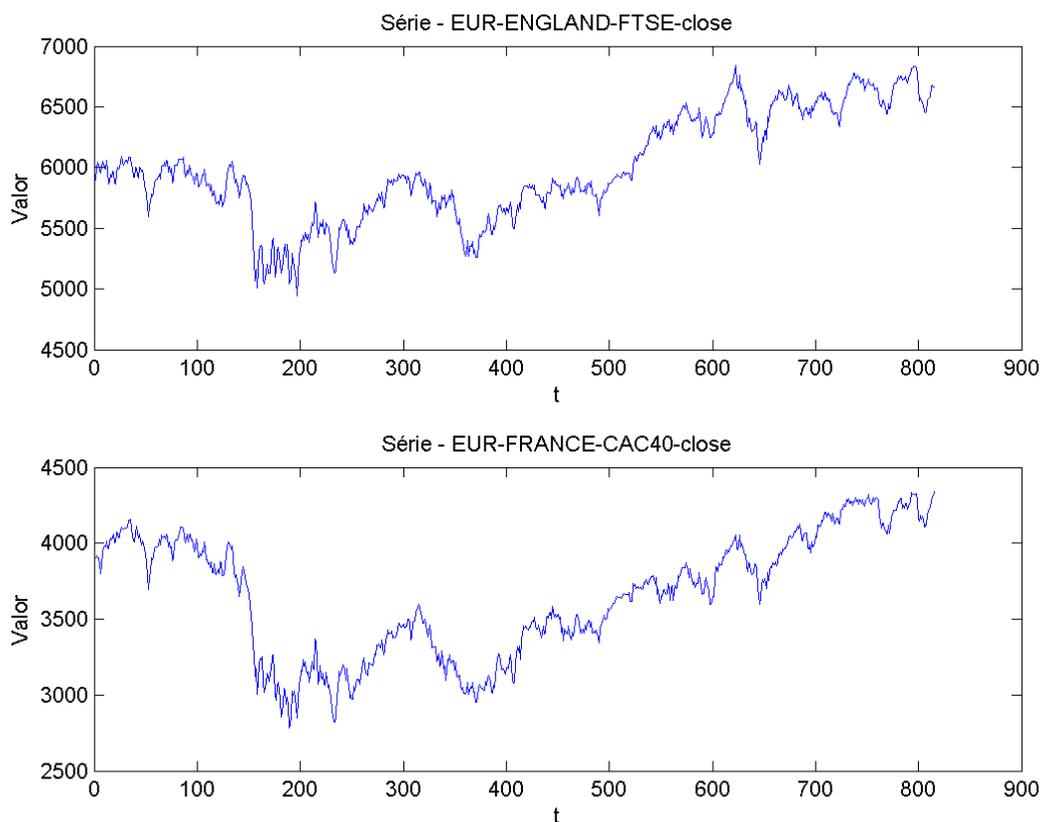


Figura 3.12: Gráficos das séries diárias dos índices de França e Alemanha

3. O Milano Italia Borsa 30 Index (MIB 30) é um índice ponderado pela capitalização das empresas listadas na bolsa de valores de Milão, na Itália. Essas empresas chegam a representar 75% do volume negociado nesse mercado.
4. O índice FTSE 100 foi desenvolvido conjuntamente pelo Financial Times e pela Bolsa de Valores de Londres, sendo lançado em 3 de janeiro de 1984. O FTSE 100 (ou *Footsie*) corresponde a uma carteira de 100 ações de maior capitalização entre as mais de 2500 ações listadas nessa bolsa.

A imagem com as séries diárias dos índices MIB 30 e FTSE 100 é exibida na figura 3.13 a seguir.

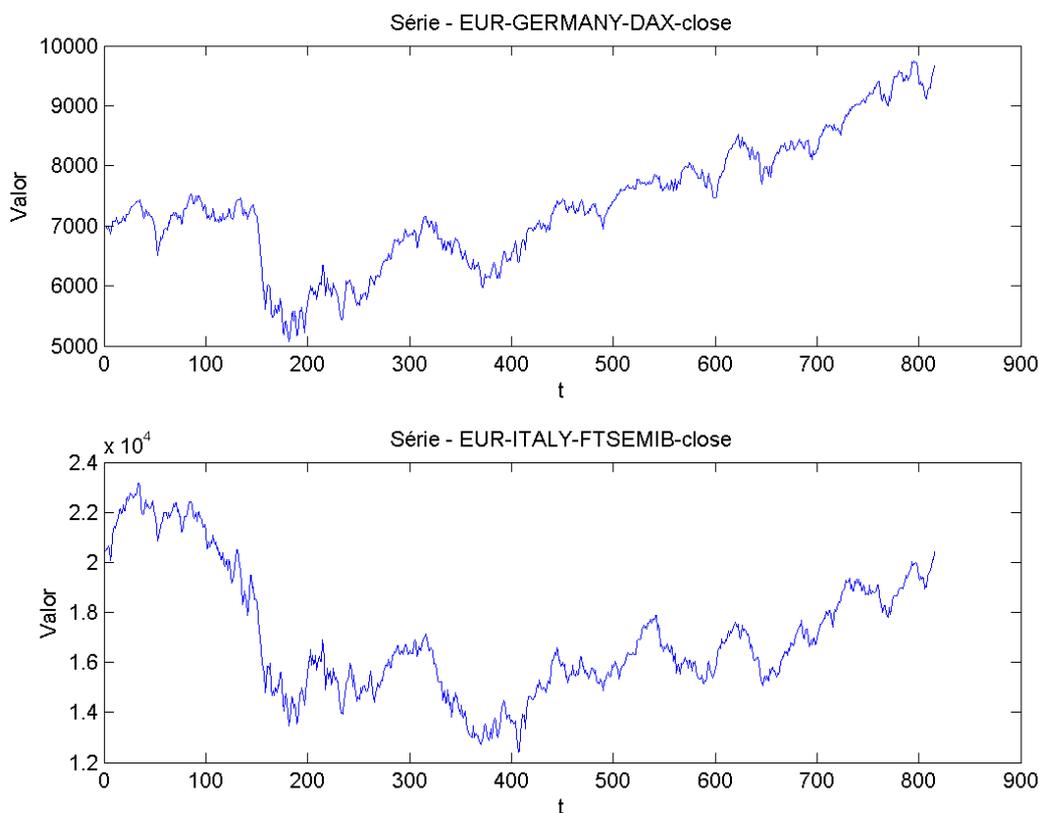


Figura 3.13: Gráficos das séries diárias dos índices de Itália e Inglaterra

Séries do Brasil

1. Os CDIs (Certificados de Depósito Interbancário) são títulos de emissão das instituições financeiras, criados como forma de captação ou aplicação de recursos excedentes. A taxa CDI mais amplamente adotada no mercado é a DI Over, publicada pela CETIP e calculada como a média das operações transacionadas num único dia, desconsiderando as operações dentro de um mesmo grupo financeiro.
2. A série obtida que representa os contratos futuros de Ibovespa, negociados na BM&FBovespa, é na verdade, uma série "rolada" a partir da junção dos valores das séries mais próximas ao vencimento, ajustadas por um contrato futuro de rolagem (que, no caso do Ibovespa, é o IR1).

A imagem com representação das duas séries descritas acima é exibida na figura 3.14 a seguir.

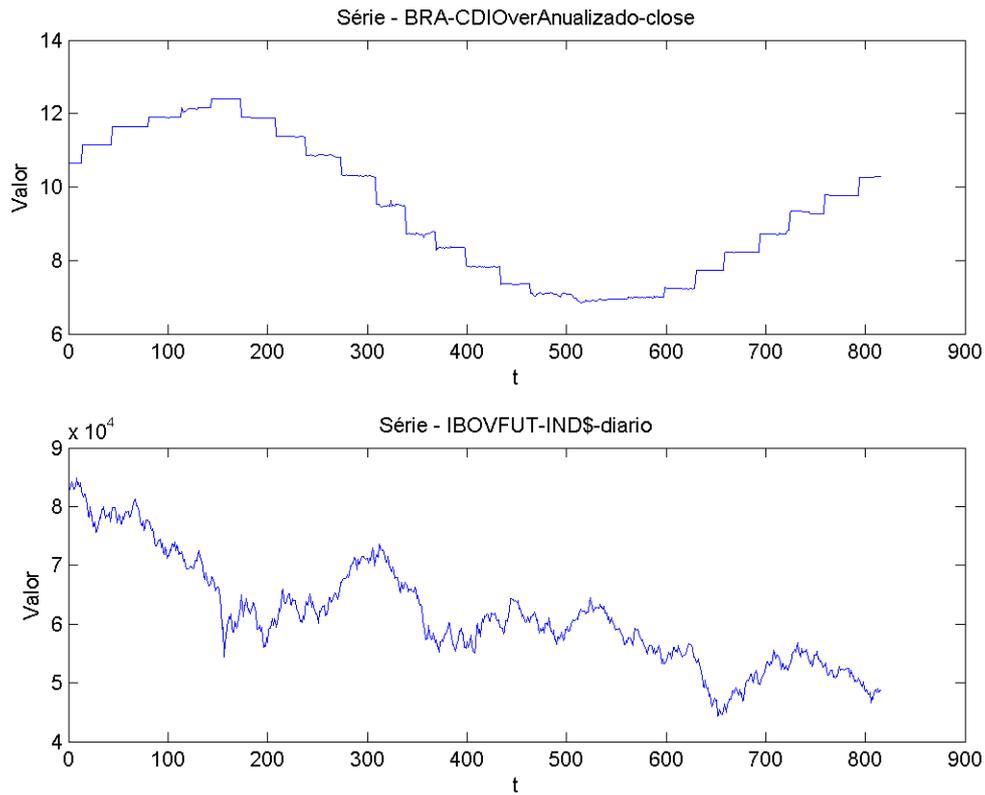


Figura 3.14: Gráficos das séries diárias do CDI e do Futuro de Ibovespa

3. O dólar PTAX é taxa de câmbio de dólar calculada ao final de cada dia, pelo Banco Central do Brasil. Basicamente, é a taxa média de todos os negócios com dólares realizados naquela data no mercado interbancário de câmbio, com liquidação em $D + 2$.
4. A série EUR/BRL representa a cotação do Euro (moeda na União Europeia) em relação ao Real brasileiro.

A figura 3.15, a seguir, exhibe os valores diários para as séries de Dólar PTAX e EUR/BRL.

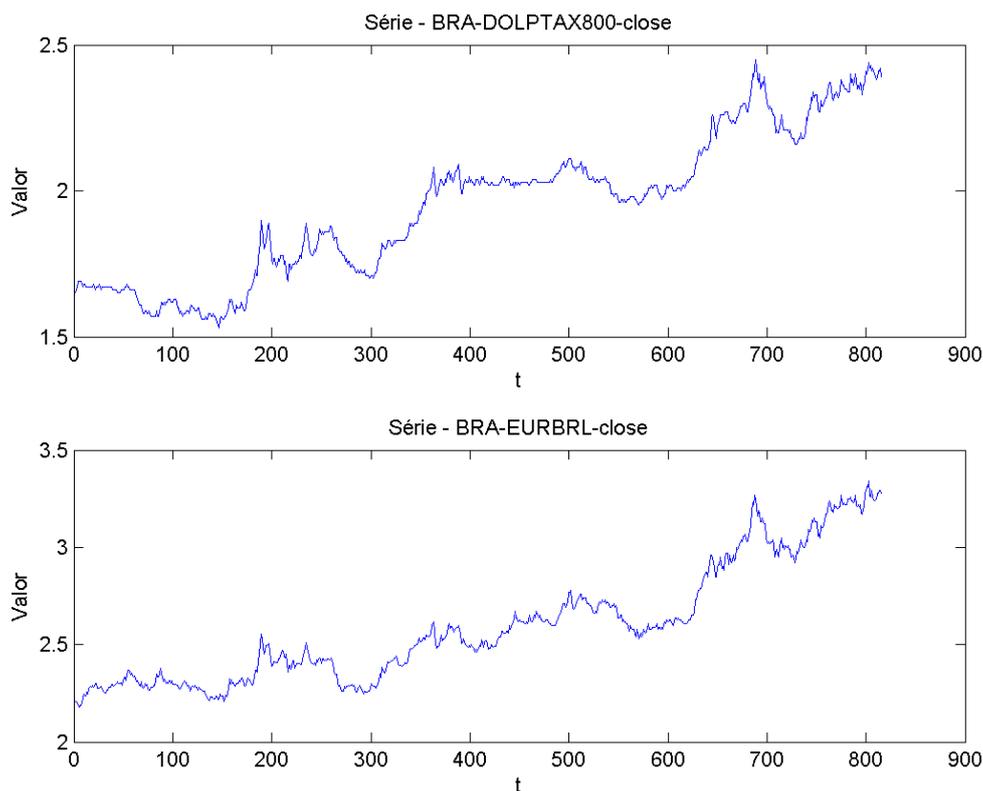


Figura 3.15: Gráficos das séries diárias de Dólar PTAX e Euro

5. O índice FGV 100 foi criado em 1986, tornando-se importante referência de desempenho das ações de segunda linha nas bolsas brasileiras [ref. SITEFGV]. Sua carteira teórica é formada por 100 papéis de 100 empresas privadas não-financeiras, negociadas na BM&FBovespa, pelos critérios de excelência empresarial, liquidez em bolsa e patrimônio líquido.
6. O Índice Dividendos (IDIV) da BM&FBovespa mede o comportamento das ações das empresas que se destacaram em termos de remuneração dos investidores, sob a forma de dividendos e juros sobre o capital próprio, e selecionadas por sua liquidez.

A figura 3.16 exibe os valores diários para as séries FGV100 e IDIV e é apresentada a seguir.

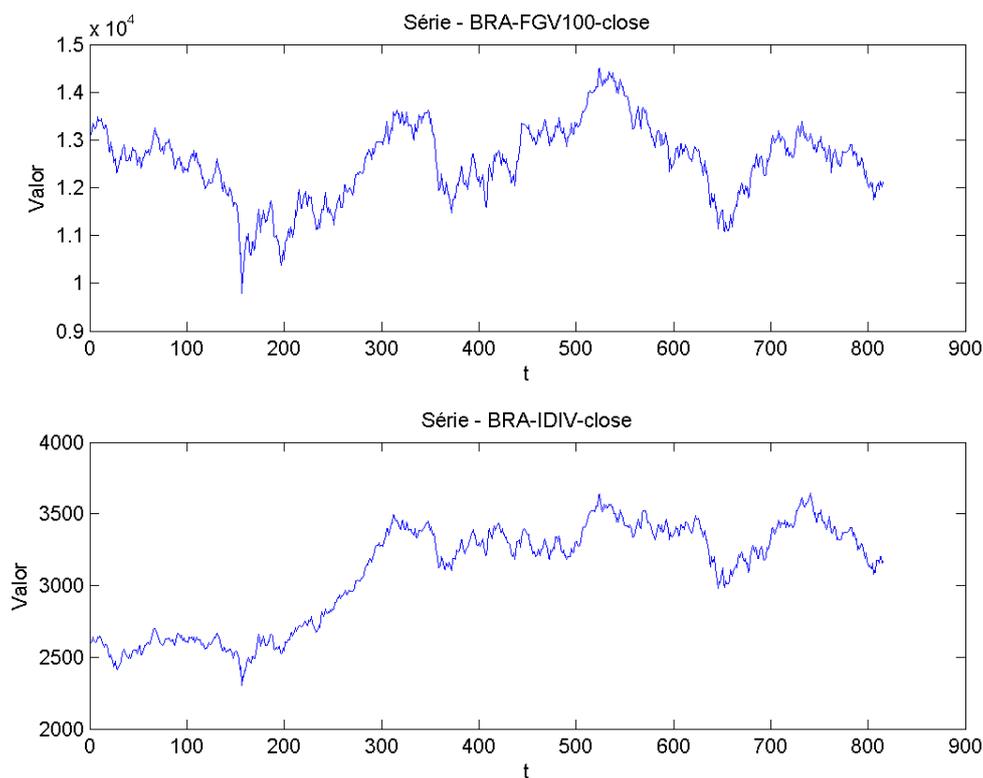


Figura 3.16: Gráficos das séries diárias de FGV100 e IDIV

7. O Índice Imobiliário (IMOB) da BM&FBovespa mede o comportamento das ações das empresas representativas dos setores da atividade imobiliária compreendidos por construção civil, intermediação imobiliária e exploração de imóveis.
8. O Índice Small Cap (SMLL) da BM&FBovespa tem por objetivo medir o retorno de uma carteira teórica composta por empresas de menor capitalização listadas na bolsa.

Por fim, a figura 3.17, exibida a seguir, contém os gráficos das séries diárias dos índices IMOB e SMLL.

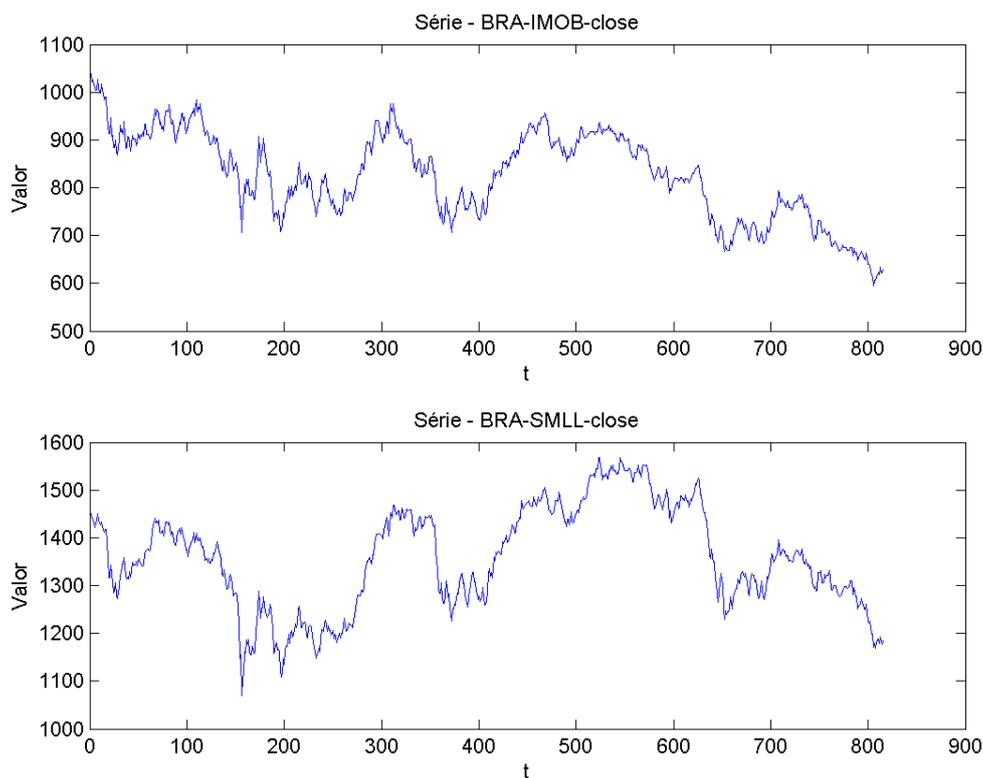


Figura 3.17: Gráficos das séries diárias dos índices IMOB e SMLL

3.2.2 Sobre a previsibilidade no mercado financeiro

O problema da predição no mercado financeiro divide pesquisadores e acadêmicos em dois grupos de opinião. Um deles acredita que o mercado é previsível e cria mecanismos e técnicas de análise para realizar essa predição. Por outro lado, há o grupo que acredita que o mercado é eficiente e qualquer nova informação que surja no mercado é absorvida por este automaticamente através de correções, logo não poderia haver predição [31].

A hipótese criada para embasar uma possível eficiência é a Hipótese dos Mercados Eficientes ou, em inglês, *Efficient Market Hypotesis* (EMH), que afirma que não há como obter maior lucro que o esperado se utilizando da predição do mercado, pois, em qualquer período, o preço de uma ação captura toda a informação disponível sobre essa ação [39]. Além disso, assim que uma nova informação surge no mercado, este absorve essa informação e se corrige, caracterizando um mercado eficiente e, portanto, não permitindo a predição [31].

Mais especificamente, há três níveis de mercado eficiente:

- Fraco: em que não é possível predizer o preço de uma ação com base na análise dos preços históricos;
- Semi-Forte: em que não é possível predizer o preço de uma ação utilizando como recurso a informação disponível no mercado;

- Forte: em que não é possível prever o preço de uma ação nem mesmo utilizando informação privilegiada (como no caso dos chamados *insiders*).

De acordo com a EMH descrita, as flutuações do mercado financeiro são baseadas no modelo aleatório ou *Random Walk*, que é representado pela equação 3.3.

$$x(t + 1) = x(t) + r(t) \quad (3.3)$$

onde $x(t)$ é o preço da ação no dia t e $r(t)$ é um valor aleatório. Com base nessa equação, a hipótese conclui que a melhor previsão para o dia seguinte $x(t + 1)$ é o valor do dia presente $x(t)$.

Analisar a EMH é importante, pois ela contradiz todas as outras formas de análise. Se superar o mercado é impossível, como afirma essa hipótese, então os métodos de análise tradicionais não deveriam resultar em uma performance melhor que um palpite aleatório [36]. Na verdade, muitos participantes do mercado conseguem superá-lo. Logo, supõe-se neste trabalho que a EMH pode não ser completamente verdadeira, na prática. Até poderia ser, se houvesse um mercado ideal e eficiente com distribuição equivalente das informações, apesar de isso não ocorrer no mercado financeiro atual.

Conforme constata Yaser [4], apesar da ampla discussão sobre a EMH, é difícil tanto prová-la quanto refutá-la. Em suas palavras, "nossa experiência pessoal nos mercados nos diz que os mercados financeiros são 'de alguma forma' previsíveis". Segundo ele, a existência de grandes períodos de tendências bem definidas e as correlações entre eventos fundamentais e econômicos que afetam os mercados são duas das muitas evidências contra a EMH.

É fácil perceber a complexidade envolvida no processo de previsão de séries temporais financeiras, sendo este um dos problemas mais desafiadores dentro de tal universo de estudo. Séries financeiras são inerentemente ruidosas, não-estacionárias e deterministicamente caótico [4]. A característica ruidosa refere-se à indisponibilidade de informações completas do comportamento passado dos mercados financeiros para capturar totalmente a dependência entre os preços futuros e passados, realizando assim uma previsão perfeita. A informação que não está incluído no modelo é considerado como ruído.

Considere o mercado de capitais como um sistema que recebe diversas informações de entrada, aqui denominada I , que I contenha todas as informações possíveis e disponíveis no instante t , sejam valores de preços, variações, indicadores, dentre outros, e que a entrada I produza uma saída y , sendo y o retorno do preço de fechamento para um determinado dia D , por exemplo.

Segundo Yaser [4], um modelo (por exemplo, uma RNA) tenta simular o mercado, mas, ao contrário de utilizar a entrada I , utiliza apenas um subconjunto das

informações necessárias, denominado por x . A informação restante não é levada em conta pelo modelo, por indisponibilidade da informação, impossibilidade de representação - como as características psicológicas dos investidores - ou limitações de processamento. Essa informação é simplesmente tratada como ruído no modelo, levando-o a obter uma performance que busca apenas uma faixa de acerto de sinal de predição (acurácia) de $(50\%+\epsilon)$, ao contrário de outras aplicações (como reconhecimento óptico de caracteres - OCR) que buscam a performance de $(100\%-\epsilon)$.

Por sua vez, a característica não-estacionária implica que a distribuição das séries temporais financeiras é variável ao longo do tempo. Por deterministicamente caóticas, quer se dizer que as séries temporais financeiras são aleatórias no curto prazo, mas no longo prazo tendem a ser determinísticas.

Uma possível solução para contornar os problemas descritos seria usar um conjunto de informações x tão grande quanto fosse necessário. Mas tal solução não é aplicável nesse caso, pois os dados de séries temporais financeiras apresentam um inerente problema de não-estacionaridade. Em outras palavras, séries temporais financeiras não possuem estatísticas constantes no tempo (como média e variância). Assim, dada a dinâmica evolução do mercado, os dados mais antigos podem ter representados padrões que já não mais se aplicam.

Outras boas referências sobre a questão da previsibilidade no mercado são encontradas em [25] e [27].

3.3 Formas de Análise em Investimentos

Uma das divisões clássicas em relação a ideologias de investimentos é a proposta por Malkiel [38], em seu livro "A Random Walk Down Wall Street", onde o autor cita duas linhas de pensamento: a teoria *Firm-foundation* e a teoria *Castle-in-the-air*.

De acordo com a teoria *Firm-foundation*, todo investimento, seja uma ação ou um imóvel, apresenta um valor intrínseco, que representa o valor real do ativo. Este valor pode ser calculado a partir de uma análise de fatores presentes e futuros do ativo. Logo, se um ativo é negociado por um preço abaixo do valor intrínseco, surge uma oportunidade de compra do mesmo. Analogamente, se um ativo é negociado por um preço acima do valor intrínseco, é descoberto um sinal de venda deste ativo.

Em contrapartida, a teoria *Castle-in-the-air* afirma que os investidores descobrem as oportunidades de compra ou venda a partir do comportamento de outros investidores, ou seja, pela psicologia do mercado. Em outras palavras, é feita uma previsão de quando e até onde os ativos irão subir ou cair, baseada no movimento da massa investidora, e então se utilizar desta análise para saber quando entrar ou sair do mercado.

Tais fundamentos, apesar de mais antigos, têm correlação com as duas princi-

pais vertentes encontradas no mercado sobre análise de investimentos, mais especificamente, sobre as duas principais escolas de análise sobre investimentos: análise técnica e análise fundamentalista.

A primeira, análise técnica, estuda a demanda e a oferta dos papéis que representam as empresas nos pregões pela análise estatística de preços e volumes negociados, avaliando as séries históricas disponíveis e criando estudos sobre seus gráficos e principais indicadores [42].

Por sua vez, a análise fundamentalista avalia o desempenho das empresas e projeta cenários econômicos futuros, verificando a capacidade de produção das empresas, a capacidade de tomada de decisão, o nicho de mercado no qual ela está inserida e seus resultados financeiros. Basicamente, os estudos são realizados a partir de seus balanços financeiros.

Nas próximas seções, serão detalhadas tais formas de análise, bem como alguns importantes indicadores de Análise Técnica relevantes para este trabalho. Por fim, é demonstrada também uma forma mais recentemente adotada por empresas do mercado financeiro, a análise quantitativa dos dados, que já possui grande aplicação e alguns casos interessantes de sucesso.

3.3.1 Análise Fundamentalista

A análise fundamentalista avalia a situação econômica, financeira e mercadológica das empresas e projeta cenários econômicos futuros, determinando, por meio de critérios, um preço justo para o ativo, também denominado valor intrínseco. Para isso, os analistas levam em consideração os chamados fundamentos da empresa, isto é, todos os fatores macro e microeconômicos que influenciam no seu desempenho.

A principal fonte de informações para os analistas é o balanço patrimonial da empresa, que no Brasil tem periodicidade de divulgação trimestral. Dentre os possíveis indicadores extraídos destas análises, destacam-se:

- P/L (Preço/Lucro)
- P/VPA (Preço/Valor Patrimonial por Ação)
- Dividend Yield (o quanto a empresa paga em dividendos, em relação ao preço da ação)
- Total de bens/Endividamento
- Lucro líquido

Além das variáveis derivadas, devem também ser levados em conta na análise as variáveis macroeconômicas, por exemplo:

- Taxa de juros básica da economia
- Índices de inflação
- Cotações do dólar e de outras moedas
- Cotações de *commodities*
- Índices de *rating* de risco

Indicadores fundamentalistas também podem se mostrar eficientes entradas para modelos de RNAs. Vários trabalhos já abordaram a questão, destacando-se o método de Anderson [20] e outros trabalhos citados em Atsalakis & Valavanis [6].

3.3.2 Análise Técnica

A análise técnica é o estudo dos movimentos do mercado, principalmente pelo uso de gráficos, com o propósito de prever futuras tendências no preço. Os movimentos podem, neste caso, ser relativos a preços, volumes ou taxas de juros relacionados ao ativo estudado.

Nela é projetado o comportamento futuro dos preços de ativos a partir de cotações passadas (séries históricas), de forma a conseguir identificar um padrão conhecido que guiará uma negociação do ativo. A análise técnica se baseia fortemente na interpretação (por muitas vezes, subjetiva) dos gráficos e, de forma menos abstrata, a partir dos indicadores extraídos das séries.

A análise técnica também pode ser vista tanto como uma forma de psicologia social, onde as linhas de tendência e outros padrões visuais ou estatísticos percebidos seriam como fotografias do comportamento dos participantes do mercado em determinado momento. Outra visão seria como forma de análise de probabilidades, baseada na ideia de que os preços se movimentam de acordo com padrões repetitivos e identificáveis, logo, os padrões tendem a se repetir de tempos em tempos.

Os indicadores aplicados ao método proposto podem ser divididos em categorias, e são os seguintes:

Médias Móveis

As médias móveis estão entre as ferramentas mais simples e populares utilizadas na análise técnica. Sua principal função é suavizar os dados da série, tornando mais fácil a identificação de tendências. Além disso, as médias móveis servem como base de cálculo para vários outros indicadores técnicos [18].

Das variadas médias móveis existentes, a mais comumente utilizada é a média móvel simples (MMS), que pode ser obtida efetuando uma simples média aritmética dos valores da série no período observado, pela equação 3.4.

$$MMS_x = \frac{\sum_{i=0}^n V_{x-i}}{n} \quad (3.4)$$

onde V_{x-i} é o valor de fechamento no dia x e n é o tamanho do período utilizado.

Além da MMS, existem outros tipos de médias que não foram utilizadas neste estudo, criadas para solucionar pequenas incapacidades da MMS. Exemplos destas médias móveis são: Média Móvel Exponencial (MME), Média Móvel Tripla (MMT), Média Móvel Ponderada (MMP) e Média Móvel Adaptativa (MMA) [47].

Bandas e Canais

As bandas e os canais são indicadores que têm relação direta com a volatilidade, podendo ajudar um analista a antecipar movimentos fortes no mercado e identificar pontos de compra e venda. Eles "delimitam" o espaço onde os preços se encontram formando as tendências ou ondas de mercado.

O exemplo mais utilizado nessa categoria é o indicador de Bandas de Bollinger, criado pelo analista técnico John Bollinger. Ele é composto por duas linhas, uma superior e outra inferior, traçadas a partir de uma determinada distância de uma média móvel, ou seja, seu cálculo se dá a partir de um cálculo de média móvel.

Outros exemplos utilizados são as Bandas de Distorção (*SkewBands*), Canais Donchian, CCI, Diferença Compra/Venda (*Bid/Offer Spread*) e Médias Móveis Envelopes [47].

Sinais e Estudos de Momento

Sinais e momentos são utilizados para revelar ao investidor os melhores momentos para se entrar ou sair do mercado, ou seja, comprar ou vender ações, através dos sinais gerados por estes indicadores.

O *Moving Average Convergence Divergence* (MACD) é um indicador de sinal simples, eficiente e que tem grande aceitação e uso por parte dos investidores. Ele utiliza duas linhas, uma chamada de rápida e outra chamada lenta, para indicar sinais de compra e de venda. Na sua forma mais comum, o MACD é usado com a linha rápida como a diferença entre as MMEs 26 e 12 dias e a linha de sinal como a MME de 9 dias da linha MACD. As equações das linhas rápida e lenta deste indicador são dadas respectivamente pelas equações 3.5 a seguir. No caso, utilizamos o histograma MACD, que nada mais é do que a diferença entre a linha rápida e a linha de sinal.

$$\begin{aligned}
MME_n(t) &= MME_n(t-1) + K * [P(t) - MME_n(t-1)] \\
K &= \frac{2}{n+1} \\
MACD &= MME_{26} - MME_{12} \\
Sinal &= MME_9(MACD) \\
HistogramaMACD &= MACD - Sinal
\end{aligned}
\tag{3.5}$$

onde $MME_n(t)$ é a média móvel exponencial de n dias, referente ao instante t , K é uma constante, $P(t)$ é o preço do ativo em t e $MME_n(MACD)$ é a MME de n dias da linha MACD.

O Índice de Força Relativa (IFR), do inglês *Relative Strength Index* (RSI), é um indicador de momento que mede, como o próprio nome diz, a força de uma ação. Ele oscila entre 0 e 100 e pode ser utilizado, numa análise, para identificar: topos e fundos, formações gráficas, suportes e resistências e, também, divergências entre o próprio indicador e preços do ativo. O IFR é calculado segundo a equação 3.6.

$$IFR = 100 - \frac{100}{1 + \frac{\sum \varphi}{\sum \omega}}
\tag{3.6}$$

onde φ é a média das variações positivas do preço da ação no período analisado e ω é a média das variações negativas do preço da ação no período analisado. Quanto menor o tamanho da janela de tempo, mais sensível será o oscilador e mais larga será sua amplitude.

Movimentos resultando em IFR abaixo de 30 são considerados como sobrevendidos e um sinal de compra é gerado quando o indicador cruza novamente a linha dos 30 para cima. Analogamente, movimentos acima de 70 são considerados como sobrecomprados e um sinal de venda é gerado quando o indicador cruza novamente a linha dos 70 para baixo.

Por sua vez, o Índice de Volatilidade Relativa (IVR) se propõe a oferecer uma medição da força existente no mercado, através do estudo da volatilidade apresentada (calculada através da evolução do desvio padrão através do tempo). A idéia é utilizar o IVR como uma confirmação para sinais de compra ou venda. Inicialmente é calculado o desvio padrão dos últimos períodos e em seguida uma média móvel desses valores.

O IVR é similar ao IFR, exceto que o IVR mede a direção da volatilidade com o desvio-padrão das variações de preços em seus cálculos, em vez de mudanças absolutas de preços, como o RSI faz. Sua forma de cálculo é dada pelas equações

3.7, a seguir. Operações de compra devem ser realizadas apenas quando IVR estiver acima de 50 e vendas, se o IVR estiver abaixo de 50.

$$\begin{aligned}
 IVR_{Down} &= \frac{MMS(p-1) + If(close > close(1), Stdev(s), 0)}{p} \\
 IVR_{Up} &= \frac{MMS(p-1) + If(close < close(1), Stdev(s), 0)}{p} \\
 IVR_{Final} &= 100 * \frac{IVR_{Up}}{IVR_{Up} + IVR_{Down}}
 \end{aligned} \tag{3.7}$$

onde p é o número de períodos e s é o número de períodos do desvio padrão.

O Momentum é simplesmente a taxa de retorno da diferença entre o preço atual e o preço N dias atrás, onde N é o parâmetro do indicador, de acordo com a fórmula 3.8 abaixo:

$$Momentum(t) = P(t) - P(t - N) \tag{3.8}$$

Além dos indicadores citados, podem ser destacados nessa categoria: SAR Parabólico, Estocástico, RSI Estocástico, Fluxo de Dinheiro e Índice de Massa [47].

Outros Indicadores

Outros indicadores desenvolvidos por especialistas para a análise técnica de ações são utilizados como complementares aos já citados, de forma a aprimorar a análise. Com destaque, podem ser citados o Saldo de Volume e a Volatilidade Histórica Anualizada.

O Saldo de Volume (*On Balance Volume*) mede a pressão de compra e venda como um indicador cumulativo que acrescenta volume nos dias de alta e subtrai volume em dias de baixa. Os grafistas costumam buscar divergências entre OBV e preços para prever os movimentos ou para confirmar as tendências de preços. Sua básica fórmula segue descrita na equação 3.9 abaixo.

$$\begin{aligned}
 Se\ retorno(t, t-1) > 0, OBV(t) &= OBV(t-1) + Volume(t) \\
 Se\ retorno(t, t-1) < 0, OBV(t) &= OBV(t-1) - Volume(t) \\
 Caso\ contrario, OBV(t) &= OBV(t-1)
 \end{aligned} \tag{3.9}$$

O OBV cresce quando o volume em dias de alta supera o volume em dias de baixa e decresce quando o volume em dias de baixa é mais forte. Um OBV crescente reflete pressão de volume positiva, que pode significar um cenário futuro com preços mais altos, sendo a recíproca também verdadeira.

A Volatilidade Histórica Anualizada é baseada na volatilidade do ativo, que é, basicamente, a medida estatística de desvio padrão, significando o quanto os preços costumam variar da média, em um determinado período. Desta forma pode-se estimar a região de maior probabilidade para as variações futuras dos preços e, com isso, saber se o mercado se apresenta com nível de variação em comportamento esperado ou atípico. A equação relativa ao desvio padrão foi previamente apresentada no capítulo 2.

Por fim, outros indicadores, também não utilizados neste estudo, mas que valem a menção [47], se encontram listados a seguir:

- **Osciladores:** Oscilador de Chaikin, Oscilador de Volume, Oscilador de Preço e Disparidade;
- **Estudo de Volatilidade:** Indicador de Volatilidade de Chaikin, Curtose e Distorção;
- **Estudo de Volume:** Volume+, Volume AMA, Volume EMA e Volume SMA, Volume Weighted Average Price (VWAP);
- **Estatística:** Desvio Padrão e Regressão Linear;
- **Estudos Adicionais:** Acumulação/Distribuição, ADX/DMI, ADXR, Aroon, Choppiness e Williams %R.

3.3.3 Gráficos das séries temporais de Análise Técnica

Esta seção foi criada com o intuito de exibir os gráficos das séries de indicadores de Análise Técnica que foram utilizados na dissertação. Os gráficos das séries de médias móveis simples, *MMS* com 10 períodos e *MMS* com 132 períodos, são exibidos na figura 3.18. Já os gráficos das séries *MACD* e *IFR* são mostrados na figura 3.19.

Continuando com o esquema de exibição em pares, a figura 3.20 mostra o *plot* das séries *IVR* e *BollingerBands*, sendo as séries *Momentum* e *OBV*, por sua vez, exibidas na figura 3.21. Por fim, na figura 3.22, é apresentada a série de *VolatilidadeAnualizada*, com período de 252 dias.

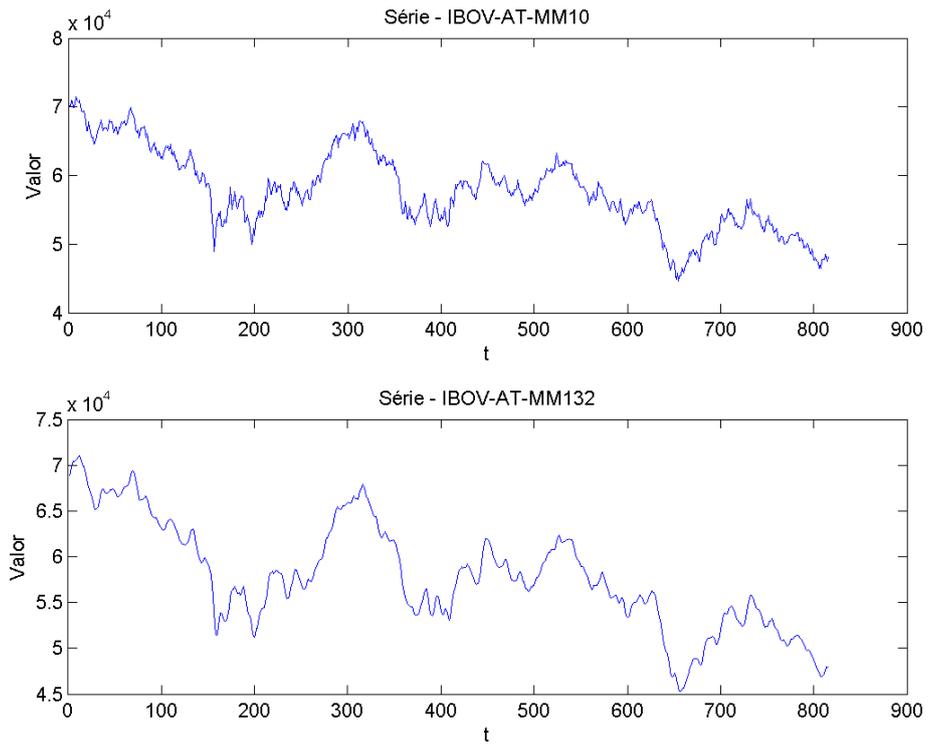


Figura 3.18: Gráficos das séries de Médias móveis

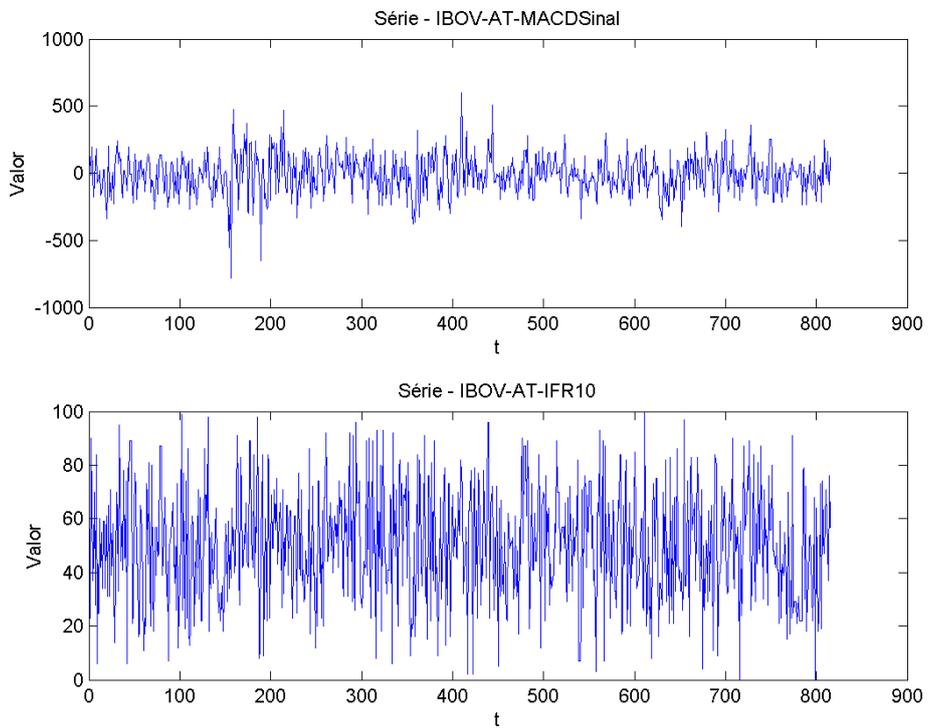


Figura 3.19: Gráficos das séries de MACD e IFR

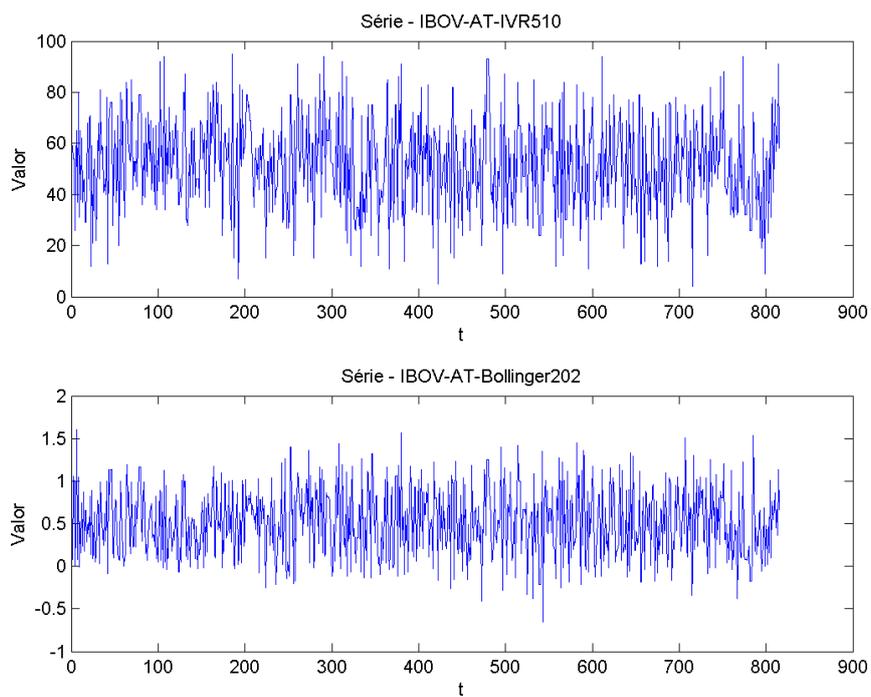


Figura 3.20: Gráficos das séries de IVR e Bollinger Bands

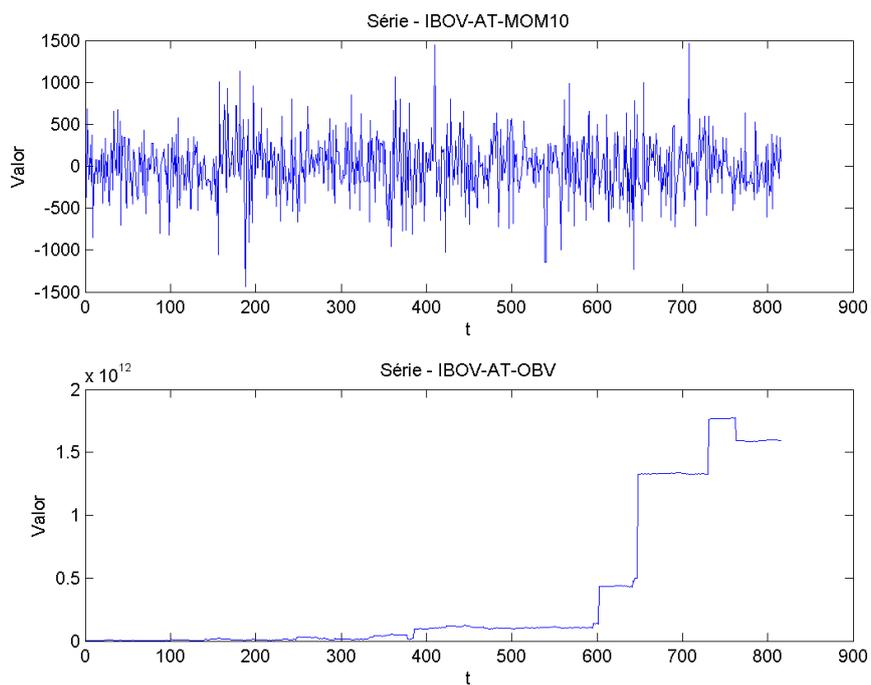


Figura 3.21: Gráficos das séries de Momentum e OBV

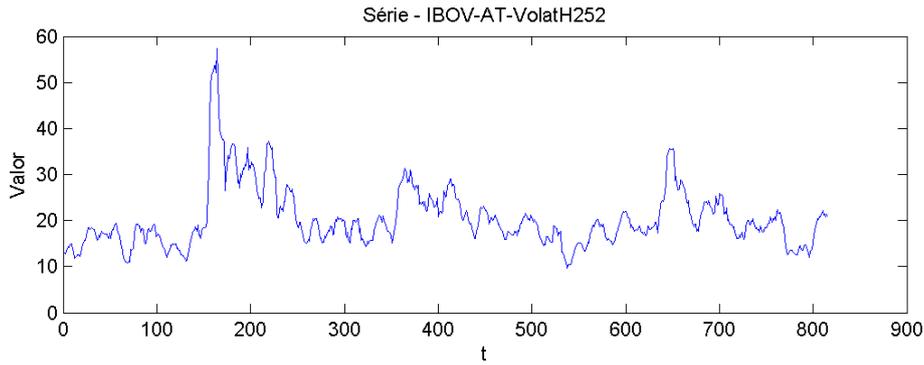


Figura 3.22: Gráficos das séries de Volatilidade Anualizada

3.4 Modelos quantitativos

Atualmente, é notável o aumento de concorrência entre as empresas do mercado financeiro que realizam e dependem de análises para operarem no mercado e que comercializam fundos de investimento. Cada empresa monta sua(s) própria(s) estratégia(s) e a disputa se evidencia sobre aquela que possui melhor infraestrutura para operar mais rápido, o que gera uma vantagem competitiva.

A competição se dá, por vezes, na otimização de algoritmos em nível de milissegundos, de forma com que sistemas de alta frequência (*High Frequency Trading*, HFT) possam superar seus concorrentes. De forma a buscar automatizar e otimizar o processo de investimento e entender o comportamento do mercado, são utilizados métodos complexos de modelagem. Não somente os sistemas que tomam decisões de operação são os explorados, como também são os sistemas de análise de risco e de gerenciamento das carteiras, que também devem dar respostas instantâneas, em certos casos.

Quando se fala em modelagem quantitativa, se envolve dois pontos principais: implementação de métodos matemáticos para precificação e investimento e execução automatizada (ou semi-automatizada) de estratégias. O universo *quant* é mais amplo que apenas esses dois pontos, mas uma discussão mais detalhada fugiria do escopo deste trabalho.

O primeiro ponto se refere às técnicas utilizadas para modelagem das séries e obtenção de pontos de operação no mercado, envolvendo não somente matemática, mas também estatística, computação e inteligência artificial - a conhecida área interdisciplinar, chamada engenharia financeira. Com isso, é crescente a busca por novos métodos e modelos que se apliquem a tarefas complexas e computacionalmente custosas no mercado financeiro. Alguns dos possíveis modelos são mencionados na listagem abaixo:

De forma direta e resumida:

- Modelos estatísticos
 - Lineares (ARIMA, GARCH, Filtros de Kalman, Suavização Exponencial)
 - Não-lineares (Teorema de Taken, Modelos de Markov, Auto-regressão linear, dentre outros) [35]
- Modelos de Inteligência Computacional
 - Métodos de aprendizado de máquina (RNA, SVM, SVR);
 - Algoritmos Evolutivos (AG, PG, Estratégias Evolutivas);
 - Lógica Fuzzy

As RNAs logo se mostram como uma poderosa ferramenta em finanças, servindo como base de diversos modelos quantitativos e surgindo como uma técnica de modelagem estatística. Elas fornecem uma ferramenta alternativa atraente para ambas as pesquisas e profissionais, pois buscam padrões que não são facilmente identificáveis pelos modelos clássicos. Em outras palavras, pode haver ineficiências no mercado e pode ser possível capturá-las usando RNAs [16].

O segundo ponto sobre modelos quantitativos envolve a criação de um sistema que emita decisões e que, de fato, opere no mercado de forma automática ou semi-automática. Esse ponto é atingido pelos sistemas de negociação, ou *Trading Systems*, explanados com seus detalhes nas próximas seções. É neles que serão ativadas as regras implementadas e serão emitidas as ordens de compra/venda ao mercado.

Considerando os pontos analisados aqui, caracteriza-se a possibilidade de se instaurar um fundo quantitativo. Um fundo quantitativo (ou fundo *quant*, como também é conhecido) é aquele gerido por *Trading Systems* e baseados em modelos matemáticos complexos de operação. Em outras palavras, esses fundos são geridos não pela tomada de decisões de um ser humano, mas sim por um modelo matemático e estatístico pré-estabelecido e executado por um computador e, em alguns casos, em altíssima frequência [55].

Uma importante constatação é que os principais nichos de interesse nesse mercado são os bancos de investimento e as gestoras independentes de ativos (*asset management companies*). Tais empresas buscam atrair cada vez mais clientes - e por consequência, valores financeiros - para seus fundos, de forma a aprimorar sua área de atuação, analisando ativos e indicadores de mercados mundiais mais rapidamente e aprimorando a receita com taxas de administração dos fundos.

3.5 Sistemas de negociação automática

Para depender cada vez menos da decisão dos analistas e traders, bem como das emoções envolvidas nas negociações de mercado, foram desenvolvidos os sistemas

automatizados de negociação, ou como são mais conhecidos, *Automated Trading Systems*, seu termo em inglês.

Os *Trading Systems* são sistemas que, dada a situação atual do mercado (cotações e indicadores), avaliam e emitem sinais de compra e venda de ativos. Além disso, eles submetem as ordens automaticamente para as bolsas de valores, necessitando pouco ou quase nada da intervenção de operadores de mesa.

Seu conhecimento é adquirido *a priori* por especialistas humanos, sendo conhecidos também na literatura por "*expert systems*". Este conhecimento é explicitado na forma de fórmulas, regras (do tipo se-então), ou outros tipos de representação sistemática. Em intervalos regulares, este conhecimento sistematizado é avaliado e ações são tomadas sobre a resposta obtida.

Bons modelos quantitativos de investimento devem passar, após sua implementação, por dois principais período de testes: o *backtesting*, quando a estratégia é efetivamente aplicada a dados históricos, e o teste de validação, quando o sistema é ligado a um sistema de corretora (*broker*) e é aplicado em tempo real, mas operando ativos virtuais. Após bons resultados nesses testes, pode ser liberado para atuação com os papéis (ativos) reais.

O uso de *Trading Systems* no contexto da predição por RNAs surge ao constatar que muitos trabalhos na literatura traduzem automaticamente bons resultados de predição em lucros no mercado. Essa inferência é incorreta na maioria dos casos, pois os puros resultados extraídos das predições não necessariamente obterão lucros reais quando utilizados por uma estratégia de investimento, sendo estes resultados completamente dependentes de uma boa estratégia para angariar lucros.

Este estudo é uma necessidade importante e que nem sempre é devidamente verificada pelos autores na literatura. Segundo Krollner [33], mais de 80% dos trabalhos (investigados em seu *survey*) relatam que o seu modelo superou o modelo de referência. No entanto, a maioria destes não consideraram as restrições presentes no mercado, como os custos de negociação e *slippage*³. Segundo ele, "um menor erro de previsão não se traduz necessariamente em aumento dos lucros financeiros", pois a avaliação deve levar em conta os custos operacionais envolvidos.

Abordando esta necessidade, foi desenvolvido um módulo de simulação que utiliza os resultados de predição obtidos pelos modelos de RNAs em estratégias de compra e venda do ativo estudado, considerando diversas variáveis presentes no mercado, como custos de corretagem, Imposto de Renda e *slippage*. Com esta implementação, pretende-se avaliar não somente indicadores clássicos (mas não menos importantes) de predição de séries temporais financeiras por RNAs, como também

³*Slippage* nada mais é que a diferença estimada entre o valor que é consultado no livro de ofertas do ativo e o valor efetivamente pago pelo determinado ativo, dada a rapidez com que uma oportunidade de operação pode ser tomada por outra corretora.

variáveis que analistas de mercado possam interpretar e, de fato, utilizar em seus próprios modelos.

O intuito aqui é mostrar o resultado da aplicação do *Trading System* que avalia suas regras em temporariedade intradiária, realizando operações sobre apenas um ativo. No mercado real, idealmente deve-se diversificar o portfólio de investimentos com outros ativos, de forma a diminuir o risco geral da carteira.

Uma observação importante, citada por Krollner [33], é que a diversificação do portfólio permite a redução do risco do emissor da empresa ou do mercado em que o índice se refere, mas a crise financeira de 2008 mostrou como podem ser devastadores os efeitos do risco sistemático de mercado sobre os retornos totais de uma carteira de investimentos. Tal risco pode ser minimizado com outros tipos de operação, como *hedge*⁴ com opções, por exemplo.

3.5.1 Testes em Trading Systems

A forma mais conhecida de testes em *Trading System* é o *backtesting*, que consiste em, dada uma estratégia, testá-la para um determinado período de dados históricos da série a ser operada. Normalmente, é realizada uma otimização dos parâmetros da estratégia considerando um período anterior ao desejado, de forma a melhor se ajustar à série temporal alvo.

O problema desta estratégia é que o mercado está em constante mudança, ou seja, seus padrões e correlações variam a cada instante. Surge a necessidade de um método mais robusto para avaliar a real performance de uma estratégia criada. Para tal, é frequentemente recomendado o *walk-forward testing*, que, de maneira sucinta, divide o período analisado em sub-intervalos e realiza a otimização de dados conhecidos para aplicação em valores desconhecidos (ou *out-of-sample*).

Logo, o esquema de janelamento utilizado no método proposto, é uma forma especializada de se realizar o *walk-forward testing*. É importante notar que o sistema de predição e o *Trading System* funcionam de maneira independente. Uma vez que as RNAs tenham concluído a tarefa de previsão, o *Trading System* é capaz de avaliar sua qualidade, realizando (ou não) operações no mercado. A simulação segue regras preestabelecidas e leva em conta as reais limitações do mercado de ações, tais como taxas de corretagem, imposto de renda e *slippage*.

A estrutura utilizada no *Trading System* implementada neste trabalho foi desenhada para atender especificamente as necessidades de obtenção dos valores de predição para aplicação de suas regras e sugere a consideração dos diversos fatores de custos e restrições impostos pelo mercado real. Portanto, as informações

⁴Operações de *hedge* (ou de proteção) têm por finalidade proteger o valor de um ativo contra uma possível redução de seu valor numa data futura ou, ainda, assegurar o preço de uma dívida a ser paga no futuro [3].

detalhadas sobre os processos e as regras estarão esclarecidas no capítulo 4, mais especificamente na seção 4.7 deste trabalho, de forma a serem inseridas no contexto da metodologia de aplicação.

Capítulo 4

Metodologia

4.1 Método criado

Com base nas teorias demonstradas e nos objetivos principais desta dissertação, foi desenvolvido um método para abordar o problema de predição de algumas das principais séries do índice Ibovespa, que neste trabalho são as consideradas séries alvo: preços de fechamento (*close*), máximo (*high*), mínimo (*low*) e a ordem em que *high* e *low* ocorrem, aqui simplesmente denominada *order*.

Tenta-se realizar as predições por meio das "séries auxiliares", ou seja, o conjunto de séries que se acredita que tenham alguma relação com as séries alvo. São consideradas séries obtidas a partir do próprio índice Ibovespa, além de outras importantes séries da BM&FBovespa, séries dos principais índices de bolsas mundiais e outros indicadores financeiros.

O *MATLAB* foi o ambiente de desenvolvimento escolhido para implementar a solução proposta. Ele apresenta funções desenvolvidas especialmente para certas tarefas de processamento de séries, operações sobre matrizes e geração facilitada de gráficos, além de possuir uma interface funcional. Isso otimiza o tempo de implementação e, em certos pontos dos módulos, há um ganho de performance sobre outras linguagens.

Para atacar o problema por uma abordagem que fosse eficiente e que se diferenciasse de outros trabalhos na área, foi definida como premissa que as principais restrições do mercado deveriam ser levadas em conta para a maior fidelidade do trabalho.

A primeira preocupação surge quanto ao processamento das séries consideradas aqui, de forma que não fossem considerados valores desconhecidos até o momento da análise. A partir de tal preocupação, aliada a um embasamento teórico sobre as vantagens de ambiente de simulação *walk-forward testing*, decidiu-se fazer uso da tradicional e útil técnica de janelas deslizantes. A utilização destas garante a correta

seleção de dados disponíveis até o momento para investigação, definindo assim, uma barreira clara entre dados *in-sample* e *out-of-sample*.

Sendo o janelamento o "esqueleto" do sistema, foram definidos dois módulos atuantes sobre os dados de cada uma das janelas: pré-processamento e aplicação de RNAs. No entorno do processamento em cada janela, aplica-se, ao fim do processo de predição, um terceiro módulo, o *Trading System*, de forma a simular, no mercado real, o poder de predição sobre os dados, além de exibir a lucratividade do método.

Uma representação gráfica explicativa sobre todo o processo pode ser conferida na figura 4.1, a seguir.

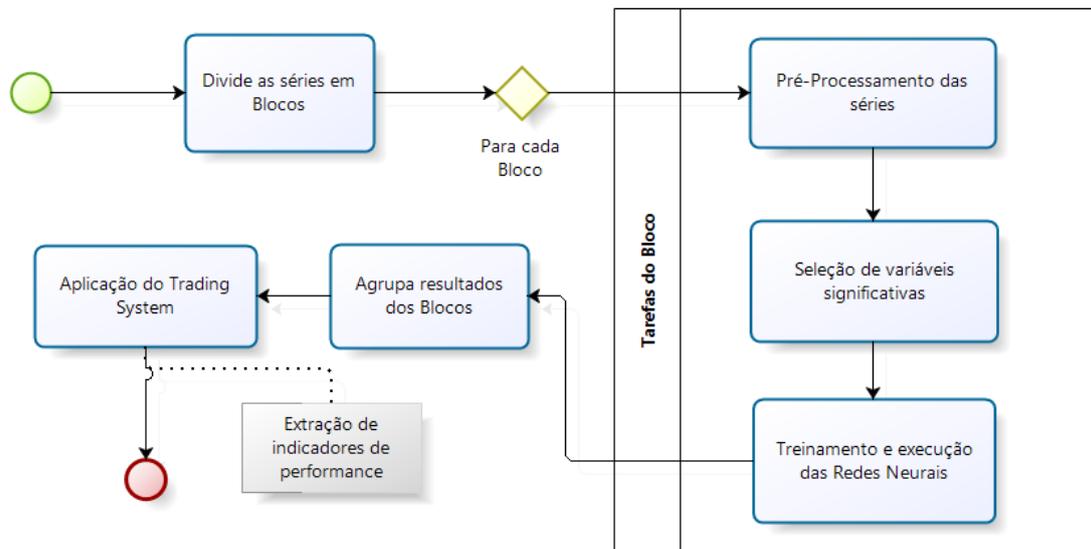


Figura 4.1: Esquema dos processos do sistema desenvolvido

Conforme discutido no capítulo 2, é parte essencial, em sistemas de previsão de séries temporais, a aplicação de pré-processamento das séries, para melhor representar os dados como entradas das RNAs. Sendo as saídas deste módulo as séries residuais selecionadas, juntamente com seus atrasos significativos, tais séries são consideradas entradas do segundo módulo, que, efetivamente, executa as RNAs e retorna as saídas estimadas para cada série alvo.

Ao término do processamento de todas as janelas, temos todos os valores estimados das séries para o período completo de testes. Neste momento, entra em ação o módulo de *Trading System*. Nele é realizado um *backtesting* de diferentes estratégias, utilizando como base os pontos estimados pelas RNAs e alguns parâmetros previamente definidos.

4.2 Aplicação de janelamento

Uma das questões mais críticas na criação de estratégias de negociação é ser capaz de saber se tal estratégia é, de fato, capaz de sobreviver às condições futuras do mercado, analisando seu comportamento em cada diferente período.

No caso desta dissertação, é aplicado um esquema de janelamento que utiliza como base a teoria descrita na seção 2.1.8. Cada bloco de investigação é subdividido em em duas partes, nomeadas bloco de Treinamento e bloco de Aplicação. O bloco de Treinamento é a seção que participa de forma integral no pré-processamento e na aplicação das RNAs, sendo, portanto, todos os dados conhecidos até momento do experimento. O bloco de Aplicação é, de forma direta, os dados *out-of-sample* de cada janela investigada. Quando na aplicação das RNAs, este bloco é dividido da forma clássica entre Treinamento/Validação/Teste, com os mesmos objetivos originais desta divisão.

Portanto, o janelamento permite que, a cada bloco, sejam definidas novas entradas e atrasos significativos, além de, é claro, treinar novas RNAs mais adaptadas ao cenário atual do mercado, conferindo ao sistema uma importante característica dinâmica. O esquema final é exibido na figura 4.2 a seguir.

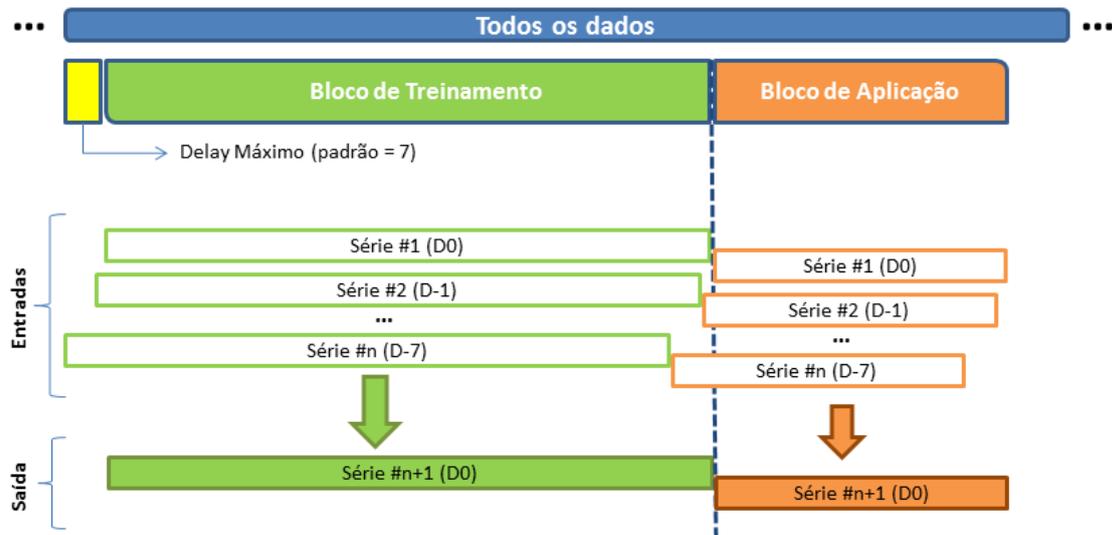


Figura 4.2: Esquema de janelamento utilizado pelo sistema

A imagem acima mostra que, a cada deslizamento da janela, são selecionados os Blocos de Treinamento e Aplicação e que há uma "sobra" de valores no início da seleção. Este segmento inicial de todas as séries temporais é incluído no pré-processamento com o objetivo de permitir a investigação dos atrasos significativo das séries, ou, em outras palavras, as próprias séries estudadas deslocadas para trás, considerando a notação: séries em $D - 1$, $D - 2$, ..., $D - N$.

O primeiro passo foi definir os parâmetros das janelas. Após diversos testes empíricos com as mais diferentes aplicações, definiu-se que serão considerados 132 dias para cada bloco de Treinamento e 10 dias para cada bloco de Aplicação, com o deslocamento da janela de 10 em 10 dias. Aplicado ao período total considerado, de 03/01/2011 até 15/08/2012, obtivemos 29 blocos para treinamento.

Uma última observação que se deve levar em conta é o descarte dos primeiros valores dos conjuntos. Por conta do parâmetro de lag máximo no pré-processamento - considerando $lag = N$ - e da transformação em séries de retorno, é necessário desconsiderar os $N + 1$ dias iniciais como valores "treináveis" no primeiro bloco de treinamento, pois estes valores serão utilizados para formação e pré-processamentos das entradas.

4.3 Obtenção dos dados

Neste trabalho consideraram-se 43 possibilidades de variáveis de entrada, conforme descrito e exemplificado com detalhes no capítulo 3, seção 3.2.1. A principal razão para a escolha de cada uma destas séries temporais é sua possível e provável relação com os movimentos do índice Ibovespa. Este conhecimento é derivado de diversas referências científicas e de experiências de investidores e *traders* do mercado.

Neste trabalho, foram incluídas variáveis de Análise Técnica (AT) - descritas e enumeradas no capítulo 3 -, além das séries já mencionadas. Como é afirmado em Gomide [26], estes indicadores são bem conhecidos por apresentar bom nível de informação para guiar *day trades* (quando ocorre abertura e fechamento de posição de um determinado ativo dentro do período de um dia).

Após a fase de escolha, foram obtidos os arquivos com dados provenientes de duas principais fontes: o *software* AE Broadcast e o *website* Yahoo! Finance. Como discorre Dantas [21], existem basicamente duas formas de se ter acesso a grandes massas de dados: ou comprando-as ou recorrendo a bases de dados gratuitas. No nosso caso, foi obtido acesso (mesmo que temporário) ao *software* pago AE Broadcast, da Agência Estado, de onde foi retirada a maioria dos dados aqui presentes, incluindo diversas séries diárias e todas as séries intradiárias.

As séries que não se encontravam no AE Broadcast, mais especificamente, Chile (IPSA-40), Canadá (S&P/TSX Composite Index), Rússia (RTSI Index) e Índia (S&P BSE Sensex), foram obtidas no Yahoo! Finance.

Foram obtidos dados diários de todas as séries compreendendo o intervalo de 03/01/2011 a 15/08/2012. Após a obtenção dos arquivos, foi necessário um esforço para tratamento dos dados contidos nestes, distinto para as séries diárias e para as intradiárias, conforme descrito a seguir.

Por sua vez, os dados do índice Ibovespa foram obtidos em periodicidade in-

tradiária por dois motivos: para que fosse possível extrair a série *order* diária e para considerar os valores de negociação no mercado, quando da aplicação do *Trading System*.

A conversão dos dados intradiários é realizada por um *script* à parte, que processa cada dia de negociação, de forma que sejam produzidas duas saídas: valores intradiários e diários. A saída intradiária é um arquivo com os pontos de negociação normalizados, ou seja, para todos os dias (que não sejam feriados) há valores definidos para todos os 28 pontos. Os 28 pontos são explicados considerando o período de negociação diário de 10:00 até 17:00, ou seja, 7 horas (420min), de 15 em 15min, o que resulta em $420\text{min} / 15\text{min} = 28$ valores.

Por sua vez, a saída diária é um arquivo com as seguintes informações:

- *Open*: *open intraday* do primeiro timestamp;
- *Close*: *last intraday* do último *timestamp* do dia;
- *High*: maior *high intraday* para um dado dia;
- *Low*: menor *low intraday* para um dado dia;
- *Volume*: soma dos volumes dos pontos no *intraday*;
- *Order*: ordem que *high* e *low* acontecem no *intraday*
 - 1 indica que o valor *high* aconteceu primeiro;
 - 1, caso o *low* tenha acontecido primeiro;
- *VespFeriadosBrasil*: dias em que é véspera de feriado no Brasil, 1 para verdadeiro e -1 para falso;
- *FeriadosEUA*: Marcação dos dias em que é feriado nos EUA, 1 para verdadeiro e -1 para falso;
- *VarDiaria* e *VarMaxima*: variáveis calculadas de acordo com a explicação no capítulo 3.

Além dos procedimentos para regularização dos dados, demonstrados acima, é necessário realizar um tratamento de feriados e dados faltantes, cuja especificação será mostrada a seguir, na seção 4.3.1.

Após os devidos tratamentos descritos, cada série pertencente à base de estudo possui um total de 423 pontos (590 dias corridos). É um valor considerável, dado que o janelamento restringe apenas os 140 valores da janela inicial de treinamento (132 do primeiro Bloco de Treinamento + 8 que serão desconsiderados por conta do pré-processamento), sendo todo o restante participante, em algum momento, de janela de aplicação.

4.3.1 Tratamento de feriados/dados faltantes

Mesmo apresentando boa consistência, os dados provenientes das fontes escolhidas podem apresentar pontos faltantes. Felizmente, para as séries brasileiras, todos os dados faltantes das séries eram, de fato, feriados. Para os outros índices mundiais, não se pode afirmar isto, pois não foi possível buscar uma fonte confiável (e grátis) com todos os feriados dos países estudados. De qualquer forma, o mesmo tratamento realizado para as séries brasileiras foi considerado para as séries de outros países, utilizando a premissa de que todos os pontos faltantes nas séries fossem feriados.

O tratamento de feriados foi efetuado incluindo pontos "sintéticos" na série nos dias de semana onde não houve negociação na bolsa, garantindo a corretude do pré-processamento de séries temporais e normalizando a representação das séries, ou seja, considerando as semanas completas com valores para seus cinco dias úteis (de segunda-feira à sexta-feira).

Em termos do janelamento, para cada bloco de Treinamento, a adição de um ponto sintético na série que represente o dia D faltante é efetuada calculando-se a média dos valores em $D - 1$ e $D + 1$, pois ambos os valores são previamente conhecidos no processo. A exceção é quando o dia faltante é o último ponto do bloco, quando o tratamento apenas replica as informações de $D - 1$.

Por outro lado, a inclusão de um determinado dia D no bloco de Aplicação é dada pela simples replicação das informações de $D - 1$, pois no dia D ainda não são conhecidos os valores de $D + 1$, mantendo a premissa da simulação sobre a não utilização de dados indisponíveis até o momento do estudo.

Este processo de complementação das séries é fundamental para a corretude dos pré-processamentos de séries temporais realizados. Acredita-se que este não implique em comportamentos estranhos para o processo de treinamento das RNAs, pois as séries modificadas aqui serão tratadas pela fase de pré-processamento. Uma de suas rotinas, a ser descrita nas próximas seções, é a atenuação de intrusos, onde possíveis efeitos indesejados pela adição de pontos sintéticos também serão tratados pelo processo.

4.3.2 Horários de bolsas

Ainda na fase de coleta de dados, foi realizada uma pesquisa para obter os horários de abertura e fechamento das bolsas dos índices participantes do estudo. O objetivo foi buscar índices de bolsas estrangeiras que tivessem seu fechamento em um horário anterior à abertura da BM&FBovespa, tornando possível assim utilizar os valores de fechamento da série em D_0 nas previsões da nossa bolsa também em D_0 .

Em outras palavras, considerando que a BM&FBovespa tenha horário de funcionamento de 10:00 às 17:00, o valor da série Hong Kong (Hang Seng Index) em D_0 ,

País	Horário local		Fuso Horário*	Horário no Brasil	
	Abertura	Fechamento		Abertura	Fechamento
Argentina	09:30	18:00	+0	09:30	18:00
Canadá	09:30	16:00	-1	10:30	17:00
Chile	09:30	16:30	-1	10:30	17:30
México	08:30	15:00	-2	10:30	17:00
EUA	10:30	17:15	-1	11:30	18:15
China	09:30	15:00	+11	22:30	04:00
Hong Kong	09:30	16:00	+11	22:30	05:00
Índia	09:00	16:05	+8	01:00	08:05
Japão	09:00	15:00	+12	21:00	03:00
Coreia do Sul	09:00	15:00	+12	21:00	03:00
Rússia	10:30	17:45	+7	03:30	10:45
França	09:00	17:30	+5	04:00	12:30
Alemanha	09:30	20:00	+5	04:30	15:00
Itália	09:05	17:25	+5	04:05	12:25
Inglaterra	08:00	16:30	+3	05:00	13:30

* Fuso Horário em relação ao Brasil. A diferença para o Brasil é dada pelo negativo do fuso.

Tabela 4.1: Tabela de horários de funcionamento das bolsas mundiais

por exemplo, pode ser utilizada para a predição do close do Ibovespa também em D_0 . Essa premissa não é verdadeira, por exemplo, para o valor do índice Merval, da Argentina, em D_0 .

A tabela com os horários de todas as bolsas estudadas é encontrada a seguir, na tabela 4.1:

Logo, os indicadores possíveis a utilizar em D_0 selecionados foram:

- *Open* (Ibovespa)
- China (SSE Composite Index)
- Hong Kong (Hang Seng Index)
- Japão (Nikkei 225 Index)
- Coreia do Sul (KOSPI Composite Index)
- Rússia (RTSI Index)

Não nos subordinamos ao fato de que as diferenças de horário entre uma bolsa e outra é variável, pela possibilidade do país adotar em certos meses do ano o horário de verão, inclusive o Brasil. Ao invés disso, selecionamos apenas as séries que apresentam uma boa diferença (mais que 3 horas) entre seu fechamento e a abertura da BM&FBovespa, dado o objetivo descrito.

4.4 Módulo de Pré-Processamento das séries

Conforme visto no capítulo 2, as séries temporais financeiras apresentam algumas características inerentes: componentes aleatórios (ruídos), não-estacionariedade e formação deterministicamente caótica [4]. De forma a atenuar o fator de não-estacionariedade das séries e aperfeiçoar o processo de treinamento e aplicação das RNAs, as séries auxiliares foram submetidas a uma sequência de pré-processamentos. Tais processos incluem:

- Transformação da série em série de retornos;
- Normalização dos dados;
- Tratamentos de séries temporais;
- Atenuação de intrusos.

O objetivo dos tratamentos é obter, se possível, uma série temporal residual que seja estacionária no tempo. A vantagem da utilização deste processo é a redução da influência de ruídos e de características de séries temporais como tendência e sazonalidade no processo de predição por RNAs, o que geralmente melhora o nível das predições.

Preservando a premissa inicial de confiabilidade do trabalho, o pré-processamento realizado sobre cada bloco de Treinamento é completo e os parâmetros obtidos nesta fase são armazenados para posterior utilização no pré-processamento dos respectivos blocos de Aplicação. De forma mais prática, a normalização de um bloco de Aplicação não deve considerar sua própria média e desvio-padrão, mas sim o resultante da aplicação do mesmo processo no bloco de Treinamento.

4.4.1 Transformação em série de retornos

Assim como o processo descrito pela seção 2.3.1, as séries originais foram transformadas em série de retorno, aplicando-se a primeira diferença logarítmica na série, dada pela equação 4.1, onde r_t é o valor da série de retorno no tempo t e x_t é o valor da série original no tempo t . Com esta transformação, a série praticamente perde seus fatores de tendência, ou seja, tornando a série mais estacionária.

$$r_t = \log\left(\frac{x_t}{x_{t-1}}\right) \quad (4.1)$$

Algumas séries apresentaram comportamento indesejado após a aplicação da transformação em série de retorno e, após investigação, verificou-se que não havia sentido em aplicar tal processo sobre as mesmas. As séries que entraram nesta

categoria são as séries bipolares (apenas com valores -1 e +1) e as que já foram obtidas como séries de retornos, conforme listado a seguir:

- *Order* (série bipolar)
- VespFeriadosBrasil
- FeriadosEUA
- VarDiaria
- VarMaxima

4.4.2 Normalização

Após a transformação em série de retornos, a série foi submetida a uma normalização, que escala os valores dos conjuntos de dados. Esse processo é fundamental para o bom condicionamento de processos numéricos de otimização, neste caso, para a RNA. A equação 4.2 a seguir mostra o escalamento aplicado:

$$x_t = \frac{1}{\sigma_{X_t}}(X_t - \mu_{X_t}) \quad \mu_{x_t} = 0 \text{ e } \sigma_x = 1 \quad (4.2)$$

4.4.3 Tratamentos de séries temporais

Apesar das séries geralmente já apresentarem um nível de estacionariedade reduzido, com a transformação em séries de retornos, foi realizado um tratamento de séries temporais que, basicamente, contemplou dois processos: a extração de tendências e a análise no domínio da frequência.

Extração de tendência

De acordo com Morettin [43], as séries de retornos financeiros "raramente apresentam tendências ou sazonalidades, com exceção eventualmente de retornos intradiários". De forma a confirmar ou refutar a hipótese, foi realizado um tratamento simples para extração de tendência linear das séries, por ser a mais utilizada [12].

O processo é efetuado conforme descrito na seção 2.3.3, onde foram obtidos os parâmetros a e b da reta de regressão e depois se extraiu a série de tendência da série original.

De fato, foram encontrados fatores muito próximos de zero para o coeficiente angular a - na casa de 10^{-2} -, o que significa que a aplicação da extração de tendência pouco ajudará no processo de treinamento. Isso é devido à transformação em séries de retorno, descrita na seção 4.4.1 [43].

Extração de sazonalidades e ciclos senoidais

A análise utilizada neste trabalho buscou identificar e remover sazonalidades e, mais especificamente, ciclos senoidais. Por isso, levaram-se em consideração duas ferramentas clássicas para tal objetivo: a autocorrelação e a análise de Fourier, conforme explicado, em detalhes, na seção 2.3.4 deste trabalho.

Devido à transformação em série de retorno realizada, não foram encontradas frequências significativas a remover das séries. Um exemplo de gráfico contendo o espectrograma e a função de autocorrelação (calculada até 50 *lags*) é exibido a seguir, na figura 4.3.

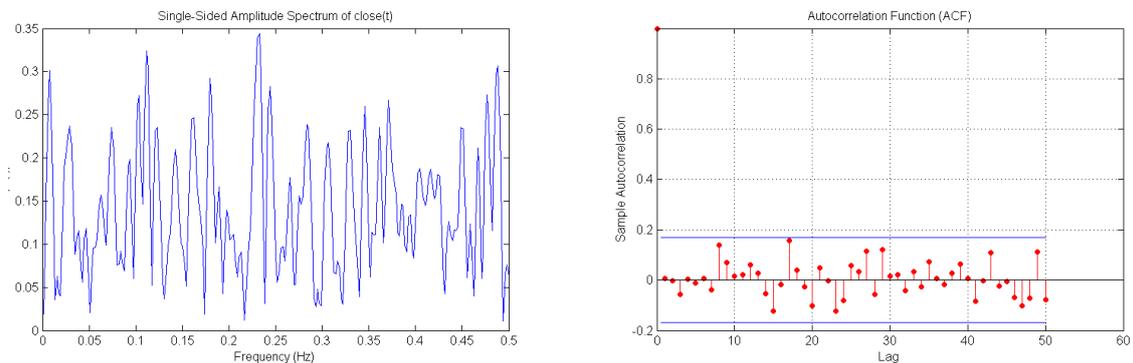


Figura 4.3: Exemplo de espectrograma e autocorrelação para a série Ibovespa *close*

Os resultados de todos os espectrogramas e funções de autocorrelação foram suprimidos do corpo da dissertação dada sua extensão e pelo fato de não terem sido encontrados ciclos senoidais em nenhuma das séries.

4.4.4 Atenuação de intrusos

Foi realizado um estudo para reduzir o efeito de valores intrusos (ou valores extremos), dada a motivação explicada na seção 2.3.5. Tais valores são caracterizados por se apresentarem acima ou abaixo de um limiar calculado a cada ponto da série, considerando valores atrasados.

Exemplificando com a série *close* completa e considerando o valor de corte como dois desvios-padrões (2σ), o total de pontos onde houve a atenuação foi de 51, aproximadamente 12% da série, o que se apresenta como um valor elevado quando comparado ao resultado obtido por Constant [10]. A figura 4.4, a seguir, mostra um exemplo de série atenuada, no caso, a série do Ibovespa *close*, destacando os pontos onde ocorreu o *clamping*.

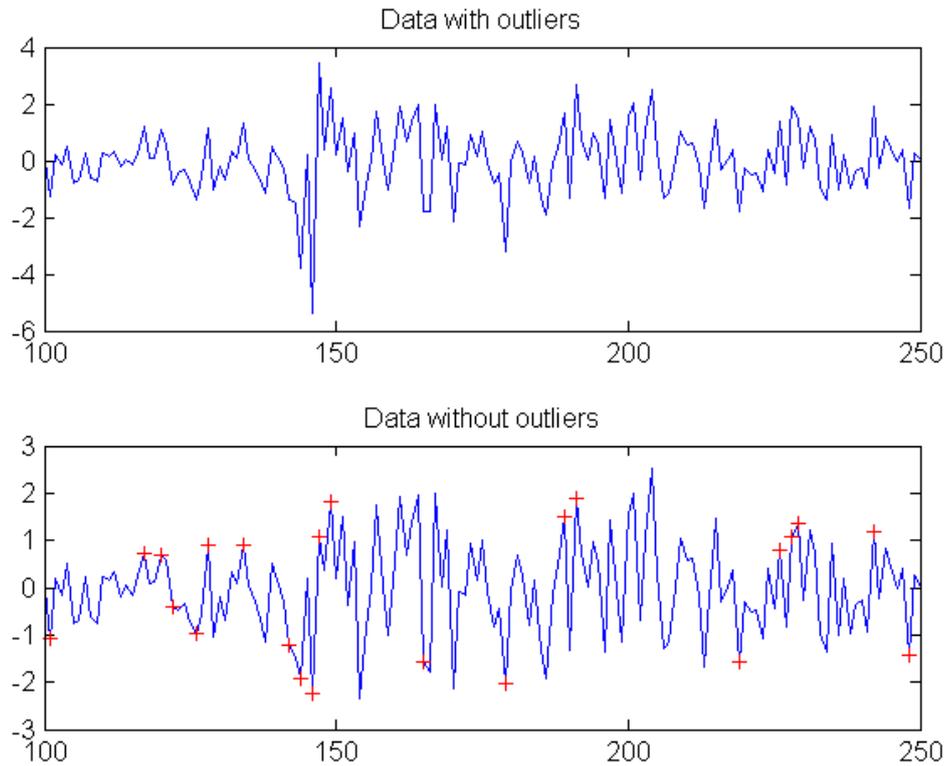


Figura 4.4: Exemplo de atenuação de outliers

É importante observar que o processo acima não se aplica às primeiras 10 observações da série, situação em que apenas se repete os 10 primeiros valores. Isso ocorre porque o processo não considera valores futuros à data em análise.

4.4.5 Estatísticas sobre os dados

Como forma de ilustrar melhor as séries que este trabalho aborda, algumas estatísticas sobre as mesmas foram obtidas ao fim de alguns passos do pré-processamento, sendo eles: antes e depois da transformação em série de retorno e após a normalização. As estatísticas extraídas estão contidas na listagem a seguir e serão exemplificadas no capítulo seguinte, de resultados.

- Média
- Mediana
- Variância
- Desvio Padrão
- Valores Máximo e Mínimo

- Obliquidade (skewness)
- Curtose

4.5 Critério de seleção de entradas significativas

Após a realização de todos os processos contidos no pré-processamento, obtemos as séries residuais. Estas séries já estão devidamente preparadas para serem utilizadas pelas RNAs que irão aproximar cada saída alvo desejada. Apesar disso, não é interessante definir todas as séries como entradas da RNA, pois desta forma, o processo apresentaria um imenso custo computacional.

Ao invés disso, foi definido um critério para escolha de algumas destas séries de entrada. Além da escolha dentre as várias séries disponíveis, foi estabelecido também um critério para seleção das séries analisadas atrasadas no tempo, com o parâmetro de máximo *lag* de atraso igual a 7, fixo para todas as séries, definido empiricamente pelos diversos experimentos realizados.

O critério aqui adotado utiliza, portanto, a informação de correlação cruzada entre a série alvo e a série analisada (que será atrasada) para avaliar a inclusão ou não da combinação de série/atraso no modelo de previsão. A correlação cruzada entre duas séries temporais $x(t)$ e $y(t)$ é dada pela correlação entre a série y tomada no instante atual, $y(t)$, e a série x tomada com um atraso de k unidades de tempo, $x(t - k)$, analogamente à autocorrelação.

Inicialmente o critério de corte de correlação significativa foi baseado na confiança de 95%, ajustando-se o ponto de corte em $\frac{2}{\sqrt{N}}$, mas o comportamento obtido não foi satisfatório, dado que para as séries *high* e *low* este esquema escolhia muitas variáveis como entrada da rede - mais de 40 em certos experimentos - e para *close* e *order*, menos variáveis que o desejado - por vezes, menos de 10 variáveis.

Por isso, preferiu-se utilizar um critério um pouco mais dinâmico, onde para *high* e *low* estreitaram-se as bandas de limite de correlação, ficando definidas em $\frac{3}{\sqrt{N}}$ e, para *close* e *order*, ampliaram-se as bandas, adotando os limites de $\frac{1.5}{\sqrt{N}}$ e $\frac{1.7}{\sqrt{N}}$, respectivamente. Estes novos limites equilibraram o número de variáveis escolhidas, que varia entre 12 e 34.

A vantagem de utilizar o esquema de janelamento, no caso desta busca por influência de séries auxiliares, é que há redução no viés da correlação variante no tempo. Quando a correlação cruzada é calculada considerando o período completo, pode ser descartada alguma relação influente para os testes *out-of-sample*, o que prejudica o resultado final do método.

Além disso, mantém-se a premissa de apenas inserir informação no modelo sobre valores ainda inexistentes no ponto de tempo investigado, pois as correlações

cruzadas atuam apenas para valores até o último ponto do bloco de Treinamento.

4.6 Módulo de Redes Neurais Artificiais

A abordagem mais utilizada na literatura de aplicação de RNAs para aproximação de funções, conforme foi revisado no capítulo 2, é a de redes *feedforward* multicamadas (do inglês *Multi-Layer Perceptron*, MLP), com algoritmo de treinamento *Backpropagation*. Dada sua simplicidade, versatilidade e capacidade de aprendizado, foi também esta a escolha para o desenvolvimento deste trabalho.

Um novo módulo foi desenvolvido contendo a implementação das RNAs, algoritmos de treinamento e toda a parte de extração de resultados, indicadores e gráficos dos processos de aprendizagem. Isso permite que, além de toda a flexibilidade de variação de parâmetros, se realize as mais distintas formas de análise que se queira.

O vetor de entradas muda de acordo com o resultado de cada bloco de Treinamento e cada saída, possuindo dimensão variável entre 12 e 34 entradas, conforme dito anteriormente. Esse é o principal motivo para que sejam realizados diversos testes para cada saída, com variação de parâmetros. Os parâmetros a serem variados - os que possuem influência direta no processo e precisam ser dependentes do tamanho da entrada da RNA - são três: taxa de aprendizado (α), número de neurônios na camada intermediária e valor máximo (em módulo) para a inicialização aleatória dos pesos sinápticos. Maiores detalhes sobre a variação de parâmetros serão expostos na seção 4.6.1, a seguir.

As camadas de saída de todas as RNA apresentam apenas um neurônio, conforme a saída desejada, pois, conforme ensaios práticos realizados em pesquisas anteriores, os resultados provenientes de RNAs com uma saída específica são melhores, comparados aos de RNAs que possuem duas ou mais saídas como alvo de predição.

Portanto, em cada bloco de Treinamento há a execução de quatro conjuntos de RNAs, uma para cada série alvo desejada (*close*, *high*, *low* e *order*). Um importante detalhe a ser mencionado é que, dada a característica bipolar da série *order*, a RNA que efetua o seu treinamento difere das demais por ser um modelo de classificação, ao invés de regressão.

A função de transferência definida para os neurônios da camada intermediária foi a tangente hiperbólica (*tgh*), sendo esta também definida para a camada de saída, no caso de RNA de classificação, e, no caso de RNA de regressão, sendo adotada a função linear $x = y$. A adoção da função *tgh* é dada pelo fato de a RNA que a utiliza ter capacidade de realizar qualquer mapeamento não-linear, de acordo com o explanado no capítulo 2. Foram também realizados testes com a inclusão de um neurônio linear na camada intermediária, além dos outros neurônios não-lineares e foi notada uma leve melhora nos níveis de erros das RNAs, de forma que esta

abordagem foi adotada como definitiva.

Para a divisão dos dados em conjuntos de treinamento, validação e teste, decidiu-se por utilizar a divisão 70%, 15% e 15%. Os 15% referentes ao conjunto de teste são sempre localizados no final da série temporal. Já os 70% e 15%, respectivamente para os conjuntos de treinamento e validação, são obtidos aleatoriamente sobre o restante da série.

Por fim, os critérios de parada do treinamento escolhidos foram: número máximo de épocas e variação mínima dos pesos. O número máximo de épocas estipulado foi de 800 e a variação mínima dos pesos, 0,0001, ambos os valores obtidos a partir da revisão de outros trabalhos e confirmados pelos testes práticos.

4.6.1 Variação de parâmetros

Dada a imprevisibilidade sobre a quantidade de variáveis selecionadas pelo critério adotado, variáveis estas que irão compor as entradas das RNAs, viu-se necessária a implementação de um esquema para variação de alguns parâmetros das RNAs, de forma a buscar sempre uma RNA que convirja para bons resultados de predição.

Para tal, alguns parâmetros das RNAs devem ser diferentes a cada bloco, são eles:

- Taxa de aprendizado (α);
- Número de neurônios na camada intermediária;
- Valor máximo para inicialização dos pesos sinápticos (em módulo).

O esquema prevê, portanto, uma combinação entre os possíveis valores destes parâmetros. Essa escolha incorre em aumento substancial de tempo de processamento, dado que a inclusão de um novo valor possível para um parâmetro acarreta automaticamente que mais RNAs deverão ser treinadas, sendo este aumento exponencial.

Logo, de forma a otimizar o tempo de computação necessitado para a execução dos experimentos, foi determinado que cada parâmetro deveria possuir apenas 3 possíveis valores, incorrendo no treinamento de 27 RNAs diferentes. Cada uma das 27 RNAs usa diferentes valores aleatórios para os conjuntos de treinamento e validação, colhendo os valores de teste do fim do conjunto.

Assim como no critério para seleção de entradas, viu-se necessário utilizar um tratamento especial para as séries close e order em relação aos possíveis valores de parâmetros. A tabela 4.2, a seguir, mostra com detalhes os valores utilizados em cada caso, sendo NE o número de entradas significativas:

A restrição necessária no caso do número de neurônios na camada intermediária é não permitir que o número de neurônios seja menor do que dois. Com o resultado das

Parâmetro	Valores - close	Valores - order	Valores - high e low
Taxa de aprendizado (α)	0.05; 0.07; 0.09	0.02; 0.03; 0.04	0.04; 0.06; 0.08
Valor máximo para inicialização dos pesos sinápticos	0.3; 0.4; 0.5	0.1; 0.2; 0.3	0.2; 0.3; 0.4
Número de neurônios na camada intermediária	NE/2-2; NE/2; NE/2+2	NE/2-2; NE/2; NE/2+2	NE/2-2; NE/2; NE/2+2

Tabela 4.2: Tabela com os possíveis valores para cada parâmetro mutável no esquema de variação

27 RNAs, cria-se um comitê e um critério para a escolha da "melhor RNA" deve ser definido. Aqui, o critério é dado por uma função sobre os erros obtidos no processo de treinamento, multiplicados por constantes de peso, resultando basicamente numa média ponderada. Os pesos de cada erro foram obtidos empiricamente, por meio dos diversos experimentos realizados tentando-se buscar qual erro deveria ter maior representatividade na função. A equação 4.3 descreve o resultado final desta função.

$$f = \|(-e1 + e2 + e3)\| \quad (4.3)$$

onde $e1$ é a média dos erros MSE em todos os conjuntos, $e2$ é a correlação entre real e estimado e $e3$ é a acurácia também entre real e estimado.

4.6.2 Avaliação do desempenho da RNA

O desempenho das RNAs treinadas foi avaliado por meio de indicadores clássicos de avaliação de RNAs, mais especificamente:

- Erro médio quadrático (MSE);
- Raiz do erro médio quadrático (RMSE);
- Erro médio absoluto percentual (MAPE);
- Correlação entre série real e prevista (*correlation*);
- Acurácia entre série real e prevista (*accuracy*).

Com a crítica durante o treinamento, é possível acompanhar a evolução dos erros de treinamento/validação/teste a cada época, para assim, determinar qual foi a época ótima. A abordagem "save the best" foi a escolhida para salvar os valores ótimos de pesos e erros e são estes valores os escolhidos no final do processo como representantes da melhor época da RNA.

Para um melhor detalhamento sobre a análise de seus resultados, foram obtidos e analisados os erros específicos para RNAs de classificação descritos na seção 2.1.7, a lembrar:

- Sensibilidade (S)
- Especificidade (E)
- Valor Preditivo Positivo (VPP)
- Valor Preditivo Negativo (VPN)
- Falsos Alarmes (FA)
- Falsas Perdas (FL)
- Erro de Classificação (e_{class})

4.6.3 Análise da qualidade das predições por faixas de retorno

Um detalhe interessante a se investigar nas predições realizadas é a capacidade de predição por faixas de variação dos retornos. Conforme observado em alguns trabalhos na literatura [10], investigou-se o fato de que, em faixas menores, o processo é mais instável do que quando lida com dados em faixas maiores. Em outras palavras, o modelo de predição é mais preciso quando os valores a serem previstos estão numa faixa maior, ou ainda, que os grandes retornos são mais facilmente acertados pelas RNAs.

Portanto, foram extraídos do processo de treinamento de RNAs os valores que melhor descrevem estas faixas. Os valores originados por este processo são, basicamente, o valor da acurácia (taxa de acerto de sinal) para cada intervalo de predição, sendo os intervalos positivos fixos e com valores $FP = 0$ até 0.1 , 0 até 0.2 , ..., 0 - até $LimiteMaximo$, similar para os intervalos negativos, $FN = -0.1$ até 0 , -0.2 até 0 , ..., $-LimiteMaximo$ até 0 . O valor da variável $LimiteMaximo$ é dado pelo limite máximo, em absoluto, de valores estimados ou reais.

Um gráfico foi gerado para auxiliar na análise destas faixas, conforme mostra a figura 4.5.

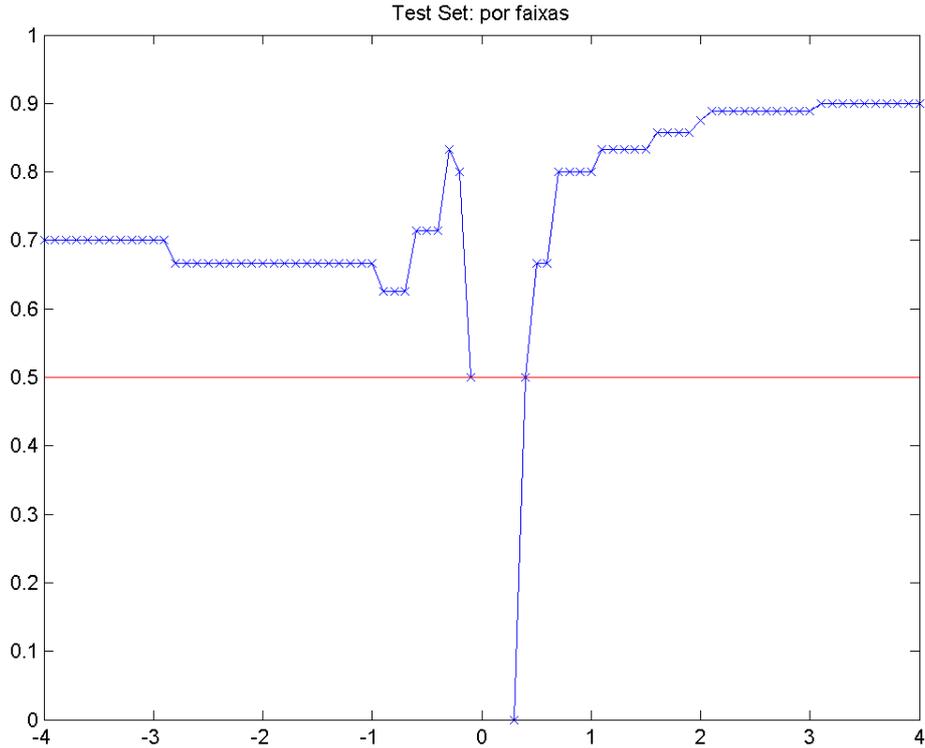


Figura 4.5: Exemplo de faixas de acurácia

Por exemplo, na imagem acima, o vetor de faixas positivas teve os valores $VP = [0.0, 0.5, 0.68, 0.68, 0.81, 0.81, 0.81, 0.81, 0.83, \dots, 0.88]$. Tais valores obtidos anteriormente significam que no intervalo entre 0 e 0.1, a acurácia foi de 0%, entre 0 e 0.2, 50%, e assim adiante, até que chega-se na acurácia do intervalo completo, 0 até `LimiteMaximo`, de 88%.

4.6.4 Estudo Comparativo

De forma a comparar o resultado das RNAs e comprovar uma possível superioridade de métodos não-lineares sobre os lineares, foi realizado um estudo comparativo com um modelo linear.

Para simular tal método linear, foi implementada uma versão específica de Rede Neural, sem camadas intermediárias e com um único neurônio linear na saída. De acordo com [51], este modelo é equivalente a uma regressão linear, ferramenta que é utilizada com certa frequência como preditor no mercado financeiro.

4.7 Módulo de Trading System

Conforme mencionado anteriormente, a simples avaliação da performance de um modelo de predição por suas saídas das RNAs é insuficiente, sendo necessário agregar valor utilizando os valores estimados como entrada de um Trading System, que, segundo Martinez [34], irá realizar a tradução dos valores estimados em decisões de negociação de ativos.

Para tal, os resultados obtidos por todas as RNAs são agrupados em uma estrutura a ser analisada pelo módulo de *Trading System*. Neste módulo, é simulada uma aplicação hipotética de uma carteira de investimentos em um determinado ativo operável no mercado. No caso desta dissertação, o ativo escolhido foi o Futuro de Ibovespa, dados o operacional de fácil compreensão e sua liquidez, conforme outros trabalhos referenciados por Krollner [33]. O objetivo deste módulo, além de realizar uma análise mais profunda sobre a qualidade das predições, é dizer se, de fato, o uso dos resultados das RNAs poderiam gerar lucros reais no mercado financeiro. Do ponto de vista prático, para um analista financeiro, estes indicadores representam a real performance do modelo.

Os experimentos com os *Trading Systems* aqui descritos levam em conta operações considerando o preço de *open* do dia (ou do *tick*, caso seja uma operação *intraday*). Essa opção, aliada ao uso de *slippage*, permite aproximar melhor o cenário real do mercado e, principalmente, permite utilizar alguns valores de variáveis em *D0*, conforme descrito anteriormente.

O *Trading System* considerado aqui pode efetuar operações de *day trading* ou *swing trading*. Também conhecido como *intraday trading*, o *day trading* ocorre quando as posições são abertas e fechadas no mesmo dia de negociação. Por sua vez, o *swing trading* ocorre quando o encerramento da posição se dá em um horizonte de curto a médio prazo, por exemplo, de dois a trinta dias após a abertura. Como será visto na seção 4.7.1, há diferentes alíquotas de Imposto de Renda (IR) incidentes em cada caso.

Além disso, é considerado aqui a livre possibilidade de compra ou venda a qualquer momento. Mais especificamente, o trading system pode efetuar operações de compra/venda ou venda a descoberto/recompra de ativos. Resumidamente, o processo analisa suas estratégias de 15 em 15min, reunindo as saídas estimadas e o preço atual do ativo, para poder decidir se efetua alguma operação (compra/venda) ou se mantém a posição atual.

4.7.1 Regras do Trading System

De acordo com Chande [13], um sistema de negociação é composto por três pilares fundamentais:

- Conjunto de regras para entrar e sair de posição de ativos;
- Mecanismo para controle de risco;
- Sistema de gestão de carteira.

Baseado em [34], deve-se ainda considerar as restrições impostas pelo mercado real, como: taxas de corretagem, taxa de *slippage* (diferença estimada entre o preço esperado e o efetivamente pago por um ativo em um trade), o volume a ser negociado e o tamanho padrão do lote, fatores estes que serão considerados como o quarto pilar fundamental.

Conjunto de regras para entrar e sair de posição de ativos

Esta é a principal parte do módulo, pois é quem dita as regras para negociação e ativa a tomada de decisão a ser efetuada pelo agente de mercado (neste caso, um agente simulado). As regras implementadas neste sistema estarão descritas em detalhes na próxima seção (4.7.2).

É importante mencionar que mais do que uma operação (compra ou venda) é permitida no mesmo dia e que a ordem em que elas ocorrem é irrelevante, pois o método considera tanto a possibilidade de compra como a de venda a descoberto na simulação.

Da mesma forma, é possível que um trade mantenha-se aberto por dias ou semanas. Se houver algum trade aberto no fim do processo, ele será encerrado efetuando-se a operação inversa, de forma a calcular os indicadores de desempenho do *Trading System* de forma correta.

Mecanismo para controle de risco

Com o propósito de controlar o risco de acentuada perda em uma única operação, foi implementado o conhecido mecanismo de *stop-loss*. Tal estratégia define uma variação mínima para o ativo (baseado no seu preço de compra), que, se ultrapassada, ativará uma operação para fechamento imediato do trade.

Após alguns experimentos variando o percentual de *stop-loss*, percebeu-se que o valor ótimo é de -7%, ou seja, caso o retorno atual de uma operação atinja -7% ou menos, ela será encerrada. Uma importante observação citada em Martinez [34] é que um *stop-loss* baixo pode evitar trades rentáveis, enquanto um alto pode causar maiores perdas.

Sistema de gestão de carteira

A gestão da carteira refere-se à forma como o dinheiro em caixa pode ser empregado em cada operação disparada pelo Trading System. Uma boa abordagem neste caso

[34] [26], é utilizar a política de aplicação de todo o dinheiro disponível em cada operação.

Assim, quando o Trading System indica uma determinada posição a tomar, todo o dinheiro disponível é investido na operação. No caso de uma compra, todo o valor em caixa (descontados os valores das restrições, conforme explanado nas próximas sub-seções) é utilizado para a operação. No caso de venda, todo o valor disponível em caixa é utilizado como "margem" para efetuar a venda a descoberto, logo, a carteira fica com um valor próximo do dobro do inicial em caixa, com a posição vendida dos ativos. Por fim, ao encerrar uma posição, todos os ativos são vendidos ou recomprados, de forma a liquidar a posição em ativos.

Apesar de todo o capital investido poder ser exposto neste tipo de gestão de carteira, o mecanismo de controle de risco escolhido (*stop-loss*) auxilia na garantia de segurança nas operações, reduzindo o risco geral da carteira.

Restrições de mercado

Para realizar uma simulação mais próxima possível do comportamento real do mercado, foram reproduzidas aqui as principais restrições e custos do mercado de ações.

Para cada operação, foi considerado um custo de R\$ 15,00 (equivalente a US\$ 6,40) fixo por operação, como taxa de corretagem. Este valor foi escolhido ao analisar o custo de corretagem fixa cobrado por diversas corretoras nacionais. É interessante notar que cada trade é composto por duas operações, uma de entrada e outra de saída, de forma que são contabilizadas duas corretagens por trade.

O impacto da corretagem pode ser reduzido ao considerar carteiras com valores iniciais cada vez maiores. Apesar disso, alguns trabalhos presentes na literatura calculam o valor final de corretagem simplesmente multiplicando o número de operações realizadas pelo custo fixo por operação, deduzindo o valor encontrado do total final da carteira. Esta abordagem pode incorrer em cenários irreais, mais especificamente quando não se considera a corretagem no momento da operação, pois pode-se sugerir investimentos com dinheiro que não existe.

Ao construir um *Trading System*, torna-se necessário considerar um fator de *slippage* para cada operação, definido pela diferença estimada entre o preço esperado e o efetivamente pago por um ativo em um trade. Essa variação ocorre no momento em que a ordem é enviada para a corretora, onde já podem existir diversas ordens de outros investidores para o mesmo ativo, e até com o mesmo preço [26]. Nestes casos, há uma priorização ordenada pelas ordens mais antigas e, com isso, o preço do ativo pode flutuar. O *slippage* é mais grave para ativos menos líquidos, pois a disparidade entre os valores de oferta e procura é maior.

Uma vez que o ativo estudado (Futuro de Ibovespa) possui significativa liquidez - média de 614.392 negócios diários -, pode-se considerar um valor baixo para o

slippage. De acordo com os testes realizados, o valor ótimo e mais coerente para este parâmetro foi de 0,01%, o que significa que os trades de compra são executados com preço adicionado de 0,01% e os de venda, com preço decrescido de 0,01%.

Ainda fazendo proveito do bom nível de liquidez do ativo estudado, o volume negociado foi de 200 ativos por trade, em média, e sempre inferior à média de negociação diária, de 614.392 negócios, o que torna o método coerente. Pelo mesmo motivo, não se torna necessário se preocupar com questões de disponibilidade para operar o ativo.

Uma vez que o mercado fracionário apresenta preços diferentes dos disponíveis nos dados obtidos, o Trading System executa somente operações no lote padrão brasileiro, ou seja, em quantidades múltiplas de 100 ações.

Finalmente, é considerado o Imposto de Renda (IR) de cada *trade*, de acordo com as regras que incidem em operações de renda variável no Brasil. A cada trade, é definido se houve lucro ou prejuízo - levando em conta a entrada e saída do trade - e este valor é somado ou diminuído de um total de incidência de imposto. Estes valores são acumulados diferentemente para *day trades* e *swing trades*, pois a alíquota de IR é diferente para cada caso (20% e 15%, respectivamente). No último dia útil do mês, é calculado o imposto devido sobre estes valores acumulados, descontando o valor devido do caixa da carteira atual.

Caso não esteja disponível dinheiro em caixa para o pagamento de IR (por exemplo, por estar em posição comprada), é retirado dinheiro de uma conta "empréstimo", que será reembolsada assim que a posição for encerrada e o dinheiro para pagamento estiver disponível.

4.7.2 Estratégias implementadas

Sendo a principal funcionalidade do módulo de *Trading System*, a aplicação da estratégia leva sempre em consideração o valor atual do ativo e as previsões necessárias, de acordo com cada tipo de estratégia. Caso nenhuma das regras seja ativada, o agente simulado não efetua operação alguma.

Três estratégias foram consideradas para comparação neste trabalho. São elas:

Levando em consideração a previsão do *close*, acrescido dos limiares α e β

Esta é a regra mais básica, considerando o estudo comparativo entre as estratégias implementadas aqui. Suas regras são definidas por:

- Caso a carteira não esteja em posição *long* (comprada) e o retorno estimado seja maior que o retorno atual (multiplicado pelo fator α), efetua compra;

- Caso a carteira não esteja em posição *short* (vendida) e o retorno estimado seja menor que o retorno atual (multiplicado pelo fator β), efetua venda.

Levando em consideração a predição do *close*, *high* e *low*, acrescido dos limiares α e β

Esta é a primeira estratégia que utiliza as predições de *high* e *low* para direcionar as operações. Suas regras são definidas por:

- Caso a carteira não esteja em posição *long* e o preço atual seja menor que o *low* estimado para o dia, efetua compra;
- Caso a carteira não esteja em posição *short* e o preço atual seja maior que o *high* estimado para o dia, efetua venda;
- Caso a carteira esteja em posição *long* e o preço atual seja maior que o *close* estimado para o dia (multiplicado pelo fator α), zera a posição;
- Caso a carteira esteja em posição *short* e o preço atual seja menor que o *close* estimado para o dia (multiplicado pelo fator β), zera a posição.

Levando em consideração a predição do *close*, *high*, *low* e *order*, acrescido dos limiares α e β

Esta regra leva em consideração as mesmas variáveis da regra anterior, com a inclusão do *order*, de forma a validar a entrada ou não no *trade*. Caso o *order* previsto seja igual a 1, mas o *low* tiver vindo antes do *high*, não é executada nenhuma operação, sendo a recíproca também verdadeira. Suas regras são definidas por:

- Caso o *order* previsto seja igual a 1, mas o *low* tiver vindo antes do *high*, não é executada nenhuma operação;
- Caso o *order* previsto seja igual a -1, mas o *high* tiver vindo antes do *low*, não é executada nenhuma operação;
- Caso a carteira não esteja em posição *long* e o preço atual seja menor que o *low* estimado para o dia, efetua compra;
- Caso a carteira não esteja em posição *short* e o preço atual for maior que o *high* estimado para o dia, efetua venda;
- Caso a carteira esteja em posição *long* e o preço atual for maior que o *close* estimado para o dia (multiplicado pelo fator α), zera a posição;

- Caso a carteira esteja em posição *short* e o preço atual for menor que o *close* estimado para o dia (multiplicado pelo fator β), zera a posição.

Seguindo a ideia de Martinez [34], a finalidade de α e β é assegurar que o trade é realizado próximo do *spread* mínimo ou máximo real, mesmo quando as previsões não são totalmente corretas. Os valores para estas variáveis foram determinadas pela execução dos experimentos, e foram definidos os fatores de $\alpha = 1.03$ e $\beta = 0.97$.

A ideia da criação destas três estratégias é realizar um estudo comparativo com a performance obtida por cada uma, de forma a avaliar o nível de aproveitamento e acurácia de cada variável de predição obtida. Comparando com todas, é incluída uma breve avaliação do retorno da estratégia com o simples Buy&Hold, ou seja, compra no início do período e venda no último dia.

4.7.3 Processamento de Carteira e Indicadores

A cada dia de processamento, é realizada uma operação especial para o processamento da carteira de investimento considerada aqui, assim como é realizada com os fundos de investimento no mercado.

Nela é calculado o patrimônio líquido (PL) da carteira, dado pela soma dos valores em caixa com o valor precificado dos ativos em posição. A marcação dos preços destes ativos é simples: o preço é dado pelo preço de fechamento do ativo no dia da análise. Com isto, é possível gerar um interessante gráfico de acompanhamento da evolução do PL, comparando com a evolução dos preços do ativo estudado.

Além do cálculo do PL, é verificada a necessidade de pagamento de IR, pagamento de empréstimo e, principalmente, são calculados todos os indicadores de performance do *Trading System*, aqui consideradas as métricas para avaliação do mesmo. São eles:

- Valor final da carteira;
- Retorno bruto / taxas / Retorno líquido (com percentuais);
- Retorno diário / mensal / anualizado (com percentuais);
- Máximo *Drawdown*;
- Índice de Sharpe;
- Número de *trades* / Número de *trades* por dia;
- Número de *trades* vencedores / perdedores (com percentuais);
- Número máximo de *trades* vencedores consecutivos / perdedores consecutivos;

- Número de *trades* em posição *long* / com lucro / com prejuízo (com percentuais);
- Número de *trades* em posição *short* / com lucro / com prejuízo (com percentuais).

Os resultados obtidos pelos experimentos, utilizando toda a metodologia descrita neste capítulo, serão exibidos com detalhes no capítulo 5.

Capítulo 5

Resultados

5.1 Experimentos

Recapitulando os capítulos anteriores, são objetivos deste trabalho: realizar as predições de *close*, *high*, *low* e *order* da série temporal do Ibovespa utilizando um esquema de janelamento e, em posse destes valores, aplicá-los a três diferentes estratégias de negociação em um ambiente simulador, o Trading System.

De forma a cumprir os objetivos, foram criados diversos subprocessos que valem a menção aqui. Como foram implementados três módulos principais (com diversas funções adjacentes), os resultados destes subprocessos serão exibidos de forma ordenada por módulo, lembrando: Pré-processamento, RNA e *Trading System*.

Será definido como "experimento geral" aquele cujos resultados são obtidos ao fim de todo este processo, com o agrupamento dos resultados de cada bloco (agregação de valores reais, valores estimados e datas). O período total compreende as datas de 03/01/2011 até 15/08/2012, resultado da agregação dos 29 blocos de treinamento do janelamento.

Foi também criada uma seção, ao fim deste capítulo, que exhibe alguns exemplos de resultados internos dos blocos, de forma a ilustrar os detalhes do processo.

5.2 Resultados - Experimento geral

Em termos de resultados do módulo de pré-processamento, há uma tabela com o número de variáveis e atrasos selecionados em cada bloco, mostrada a seguir, na tabela 5.1. Ela confirma que o processo de seleção de variáveis teve uma variação de 12 a 34 entradas, conforme destacado em negrito na mesma. Esta disparidade de valores sofreu diversas tentativas de atenuação, por meio do aumento/diminuição das bandas de corte, sendo os melhores resultados - ou seja, os que apresentaram a menor variabilidade de entradas para cada série alvo - exibidos a seguir.

Bloco	Close	High	Low	Order	Bloco	Close	High	Low	Order
1	22	22	19	23	16	29	26	24	27
2	21	24	20	18	17	29	30	27	27
3	33	28	21	13	18	32	30	28	29
4	29	28	23	23	19	32	31	26	23
5	33	29	28	18	20	21	20	27	24
6	28	32	32	12	21	25	27	27	24
7	28	32	33	12	22	29	28	29	24
8	26	30	28	14	23	28	30	25	30
9	33	30	29	16	24	34	27	26	21
10	29	30	29	29	25	30	25	22	30
11	27	31	29	21	26	28	25	22	27
12	34	32	29	33	27	30	26	23	30
13	33	32	29	22	28	28	27	24	30
14	26	32	30	26	29	25	28	29	19
15	27	32	30	25					

Tabela 5.1: Tabela com o número de variáveis significativas escolhidas para cada bloco de treinamento

	RNA				Linear			
	RMSE	MAPE	Correlação	Acurácia	RMSE	MAPE	Correlação	Acurácia
Close	0,8081	0,0048	0,4813	0,6015	0,9932	0,0133	0,3606	0,5654
High	0,7093	0,0028	0,8048	0,7985	0,8114	0,0092	0,6226	0,7315
Low	0,6807	0,0030	0,7423	0,7682	0,8210	0,0101	0,5476	0,7244
Order	0,9356	0,0037	0,3559	0,6621	0,9477	0,0142	0,2943	0,5512

Tabela 5.2: Tabela com os resultados obtidos para os modelos, comparando a utilização de modelos com RNA e com RNA Linear

Por sua vez, os resultados obtidos pelas RNAs são exibidos na tabela 5.2, que agrupa os principais indicadores extraídos da performance final dos valores *out-of-sample* previstos. Tais resultados englobam: os erros RMSE e MAPE, a correlação entre reais e previstos e a acurácia (acerto de sentido) entre, novamente, reais e previstos.

De forma a implementar um estudo comparativo, foi utilizado um modelo linear, simulado pela RNA com apenas uma camada, com um único neurônio linear. Conforme visto no capítulo 2, este modelo se equipara a um regressor linear e seus resultados são exibidos logo após cada resultado de RNA [51].

Em termos de gráficos, para o módulo de RNA, foram obtidos os itens listados a seguir, juntamente com seus objetivos.

- Comparação real versus estimado: valores estimados no eixo x e valores reais no eixo y . Tem como objetivo comparar a qualidade da predição, pois quanto

mais próximo da reta $x = y$ (plotada em vermelho), menor o erro de predição;

- Comparação real e estimado versus tempo: tempo total do experimento no eixo x e valores reais (em azul) e valores estimados (em vermelho) no eixo y . Tem como objetivo verificar a aproximação da curva estimada em relação à real, verificando também com qual frequência há acerto de sentido;
- Faixas de acurácia: conforme explicado na seção 4.6.4, o eixo x contém as faixas de predição, $FP = 0$ até 0.1, 0 até 0.2, ..., 0 até LimiteMaximo e $FN = -0.1$ até 0, -0.2 até 0, ..., -LimiteMaximo até 0. No eixo y , os valores determinam a porcentagem de acerto de sinal em cada faixa descrita aqui.

Os gráficos com os resultados finais para as séries alvo serão exibidos nas seções a seguir.

5.2.1 Resultados para as séries alvo

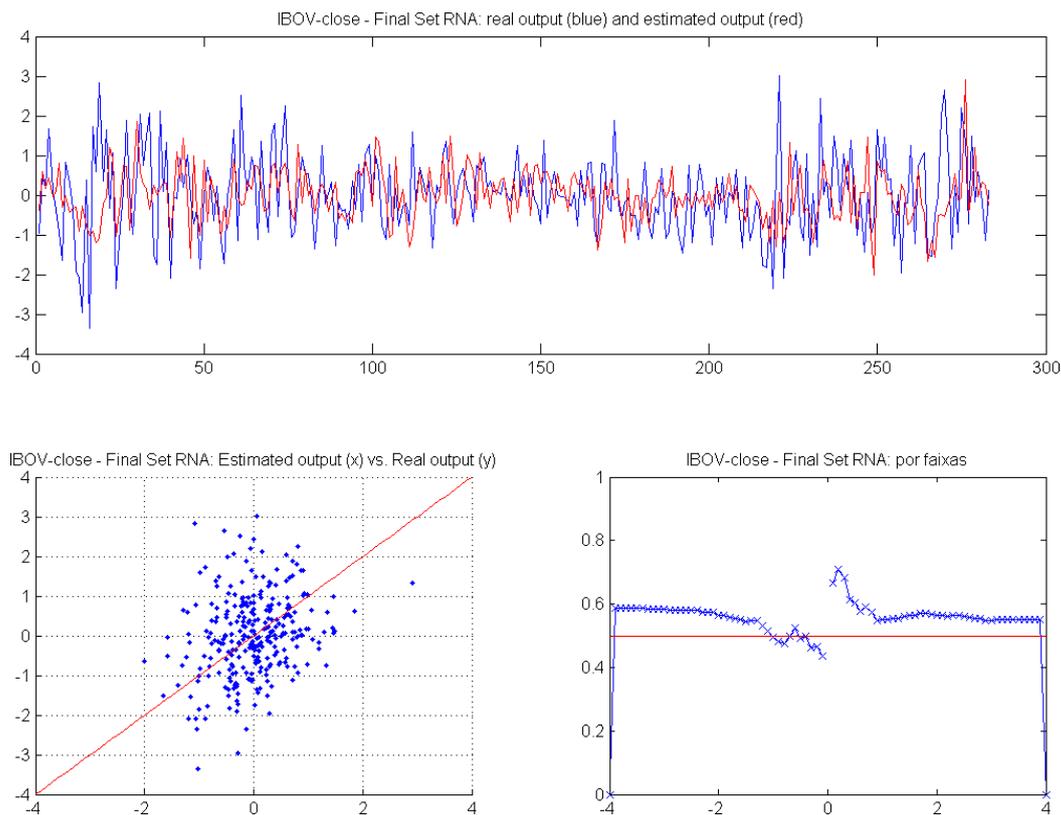


Figura 5.1: Resultados finais - Modelo RNA - Ibovespa *close*

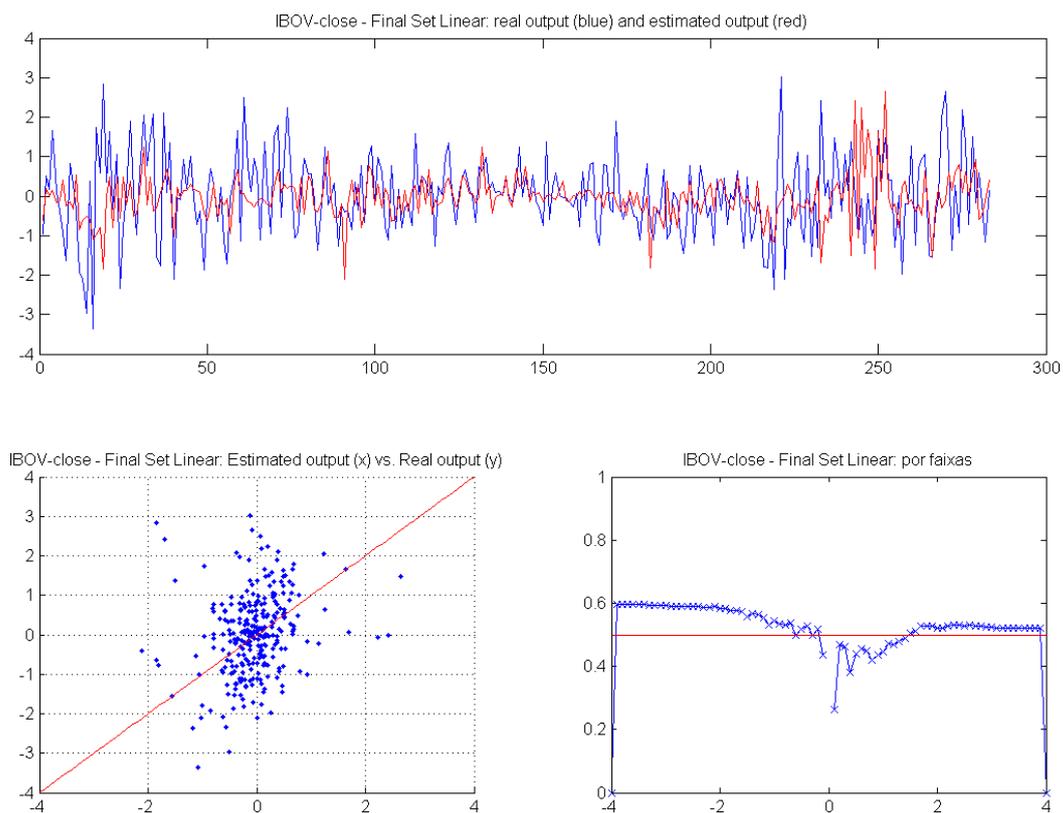


Figura 5.2: Resultados finais - Modelo Linear - Ibovespa *close*

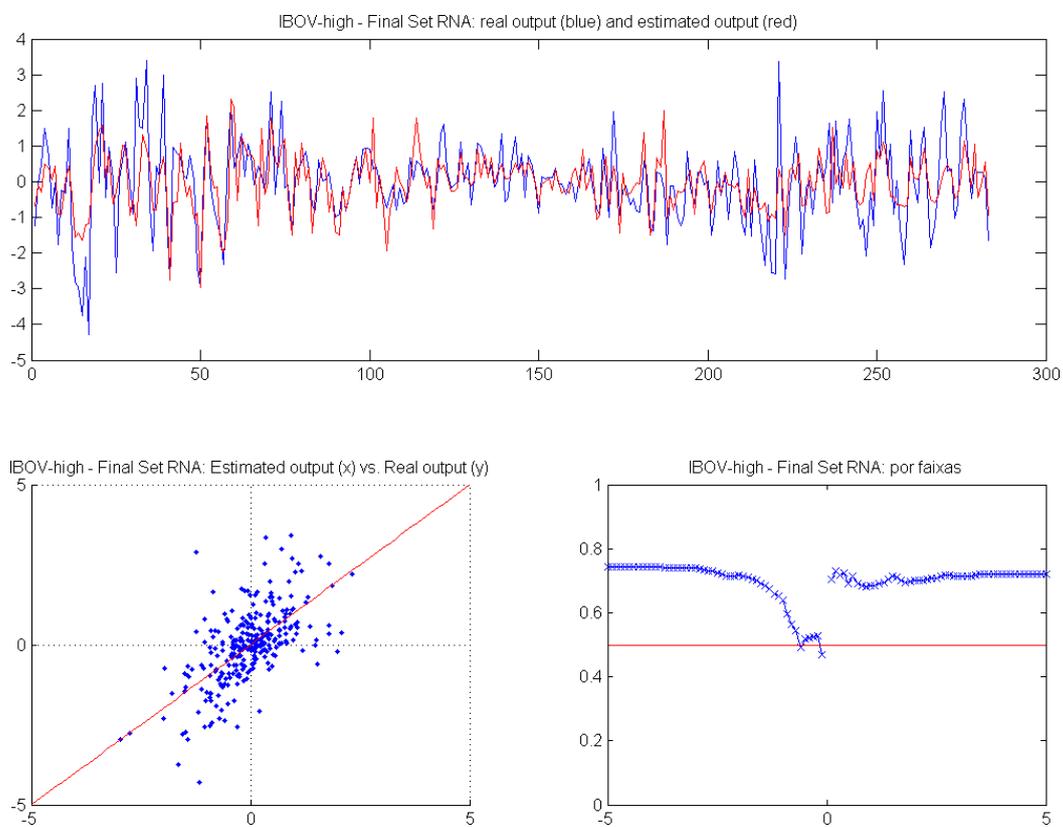


Figura 5.3: Resultados finais - Modelo RNA - Ibovespa *high*

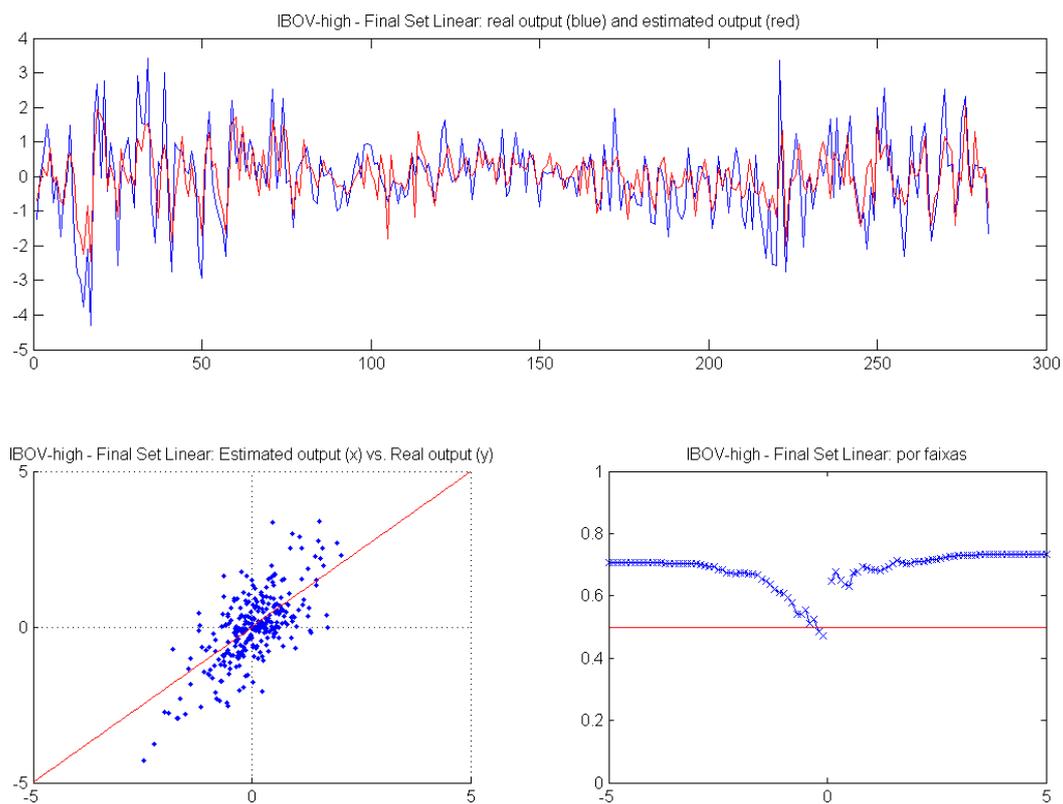


Figura 5.4: Resultados finais - Modelo Linear - Ibovespa *high*

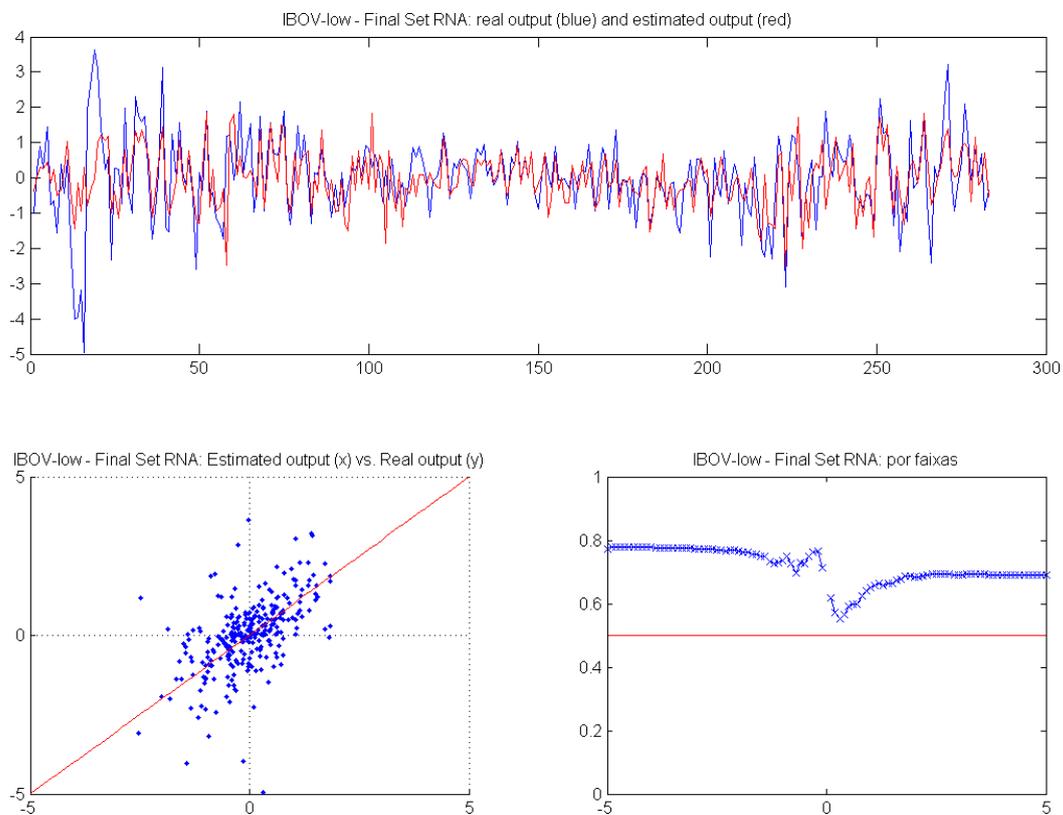


Figura 5.5: Resultados finais - Modelo RNA - Ibovespa *low*

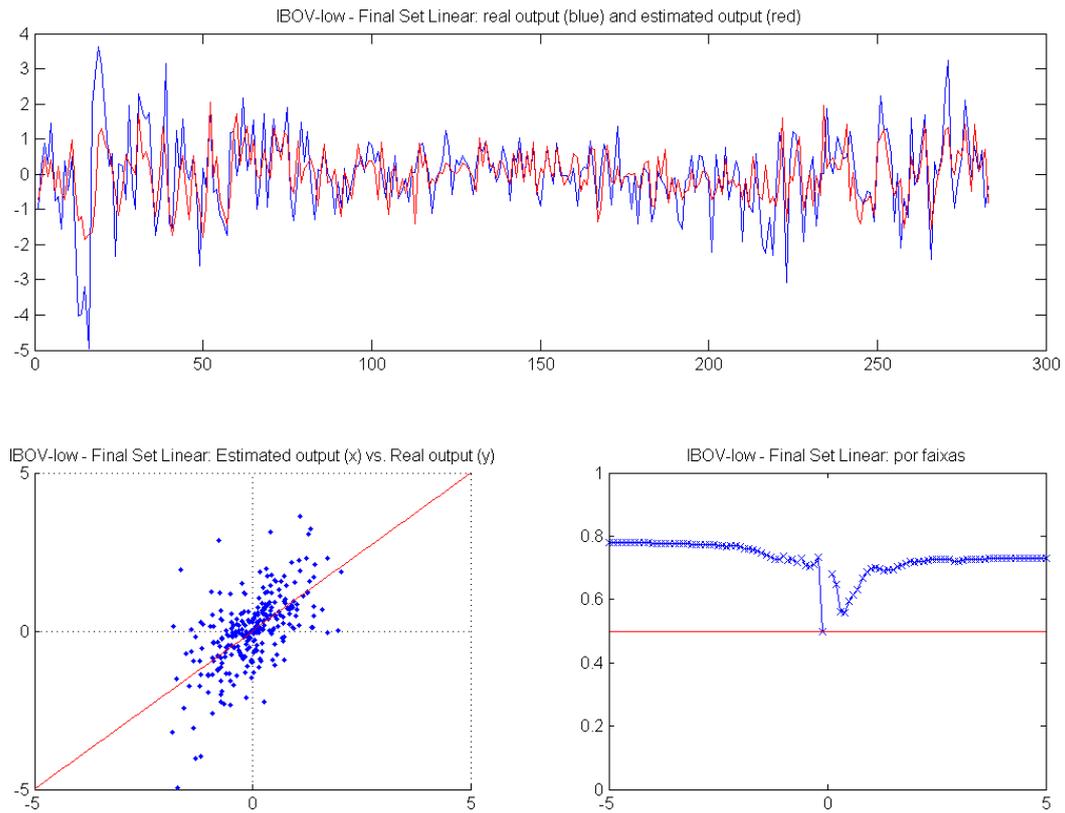


Figura 5.6: Resultados finais - Modelo Linear - Ibovespa *low*

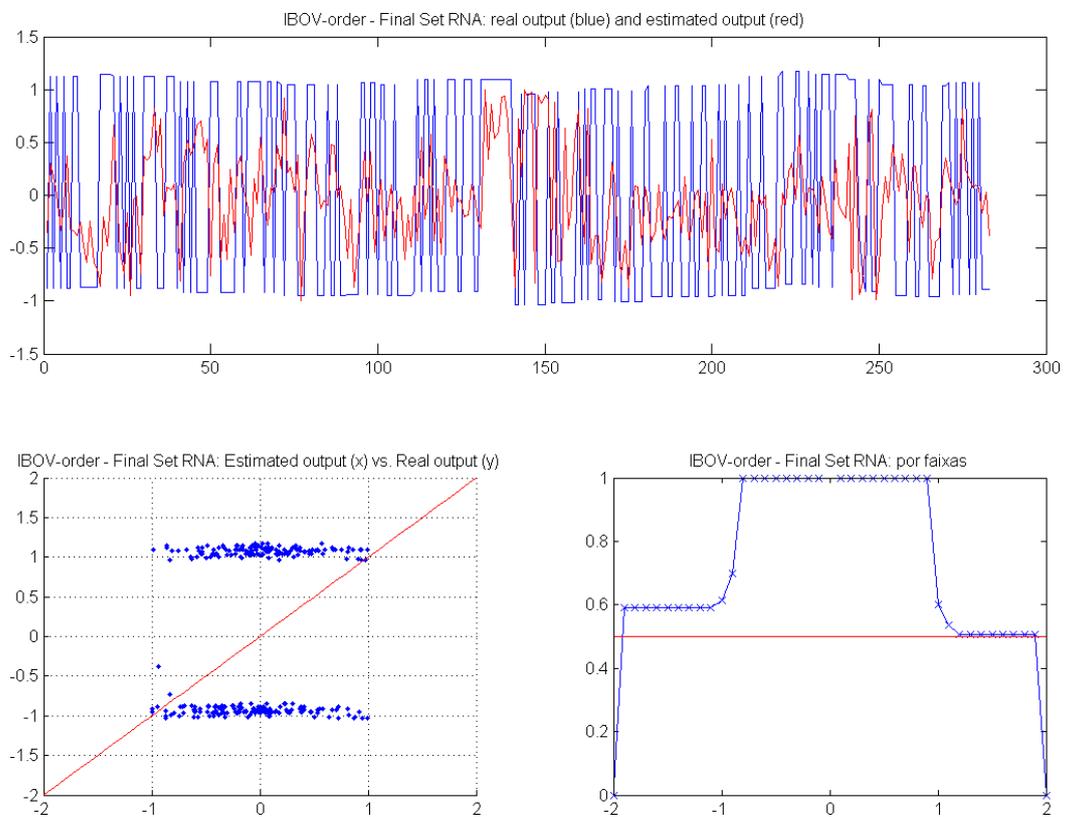


Figura 5.7: Resultados finais - Modelo RNA - Ibovespa *order*

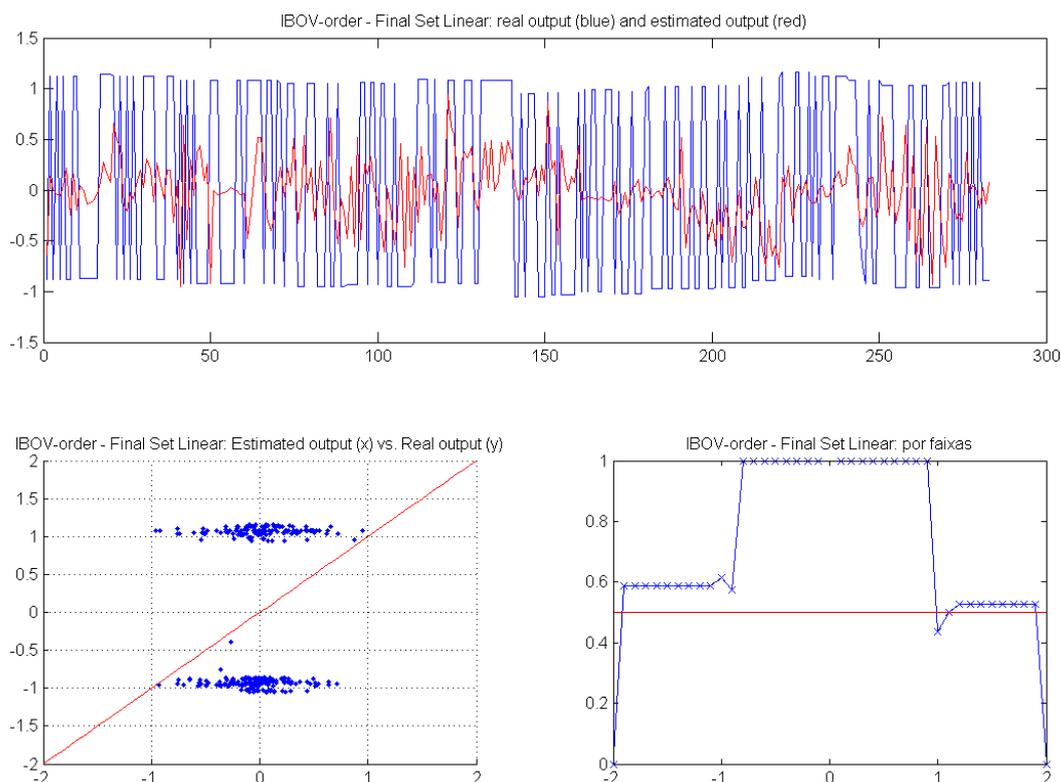


Figura 5.8: Resultados finais - Modelo Linear - Ibovespa *order*

Algumas observações podem ser obtidas sobre os gráficos exibidos. A primeira é a diferença da forma do *scatter plot* real *versus* estimado para a saída *close*, que no modelo que utiliza RNAs tem uma compactação mais bem definida e uma maior tendência a se aproximar da reta $x = y$ plotada em conjunto.

A segunda observação diz respeito à dificuldade de predição do *order*, em ambos os modelos. Os gráficos que comparam os valores reais e previstos confirmam tal afirmação, vista a dificuldade para a curva estimada se ajustar à curva real. Isso também é confirmado pelos resultados obtidos na tabela 5.2, onde a acurácia foi apenas de 55% para o modelo linear, melhorando um pouco para o modelo de RNAs.

Nota-se também que gráfico de faixas de acurácia do *order* tem uma queda acentuada do índice de acerto quando chega nas faixas de 0 a 1 e de -1 a 0. Tal comportamento é explicado pelo fato de suas faixas de valores reais serem próximas, dado que seus valores positivos variam de 0,96 a 1,17 e os negativos, -1,03 a -0,7, explicando a queda brusca.

5.2.2 Resultados para Ibovespa *order* - Classificação

Para a RNA da saída alvo *order*, foram ainda obtidos resultados específicos de classificação. A seguir, são exibidos estes valores extraídos em forma tabular (5.3) e nas figuras 5.9 e 5.10, a comparação dos histogramas de modelos RNA e Linear, indicando visualmente as estatísticas de falsos/verdadeiros positivos e falsos/verdadeiros

	RNA	Linear
Sensibilidade (S)	0,6027	0,5271
Especificidade (E)	0,6241	0,5714
Valor Preditivo Positivo (VPP)	0,6647	0,5075
Valor Preditivo Negativo (VPN)	0,5563	0,5906
Falsos Alarmes (FA)	0,3759	0,4286
Falsas Perdas (FL)	0,3973	0,4729
Erro de Classificação (e_{class})	0,2938	0,2827

Tabela 5.3: Resultados finais de classificação - Ibovespa *order*

negativos.

Nota-se que na maioria dos indicadores, a RNA superou o modelo linear, apenas uma pequena desvantagem no Valor Preditivo Negativo (VPN). No gráfico, foi observado que as distribuições são mais próximas da forma da distribuição normal para o modelo linear, significando que a RNA conseguiu identificar melhor os valores mais afastados de zero, tanto positivamente quanto negativamente.

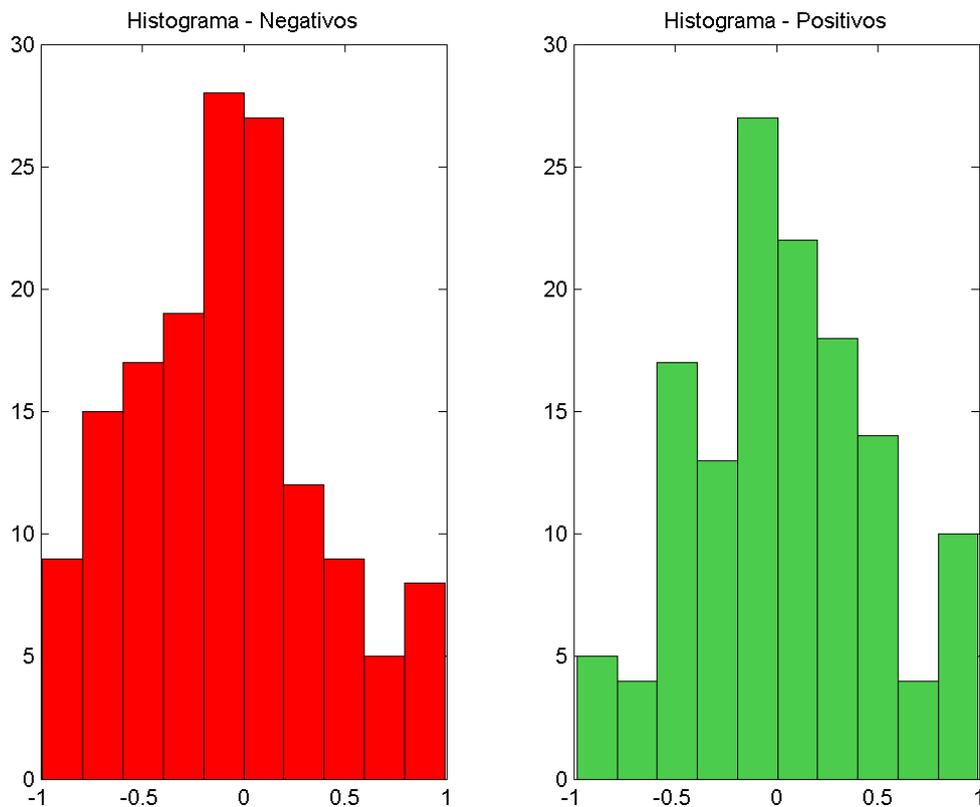


Figura 5.9: Histograma de classificação - Modelo RNA - Ibovespa *order*

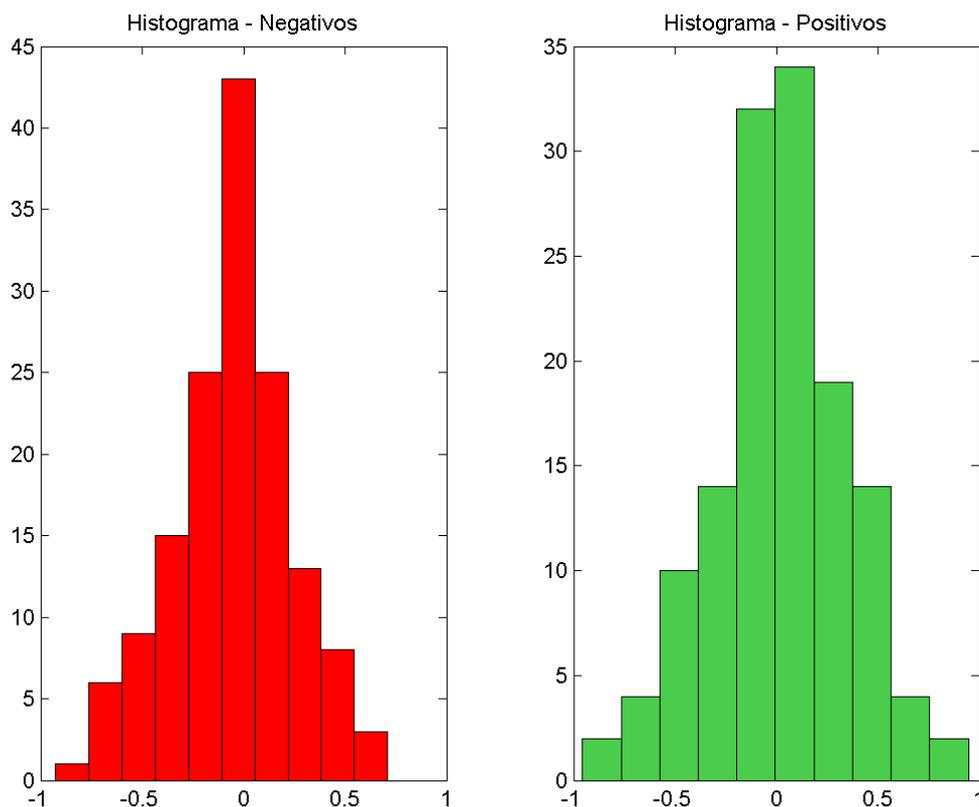


Figura 5.10: Histograma de classificação - Modelo Linear - Ibovespa *order*

5.2.3 Resultados de Trading Systems

Para o módulo de *Trading System*, foram comparadas as três estratégias implementadas, a lembrar: estratégia 1 (baseada na predição do *close*), estratégia 2 (baseada nas predições de *close*, *high* e *low*) e estratégia 3 (baseada nas predições de *close*, *high*, *low* e *order*).

Para cada estratégia, ainda, foram considerados três tempos distintos de janela de análise, mais especificamente: diário (na abertura da bolsa, horário de 10:00), de 1 em 1h e de 15 em 15min. A tabela 5.4 comparativa dos resultados obtidos é mostrada a seguir. Os valores de α e β são respectivamente 1,03 e 0,97, o *slippage* considerado foi de 0,1%, dado o considerável volume de negociação do Ibovespa Futuro.

Nota-se que a estratégia 2, aplicada apenas na abertura da bolsa é a que obteve o melhor resultado de retorno financeiro, superando as outras estratégias e o simples *Buy & Hold*¹, que obteve um retorno de -17,05%. Uma observação que se nota na tabela descrita é que quanto mais trades, menor o retorno obtido pelas estratégias. Outro ponto a comentar é que se esperava uma melhora dos resultados ao utilizar a predição do *order*, mas isso não ocorreu, provavelmente pela questão restritiva que tal variável apresenta, possivelmente impedindo trades de serem finalizados em

¹*Buy & Hold* é a estratégia de compra no início e manutenção até o fim do período de análise.

	Estratégia 1 - Close						
	Valor Final Carteira	Ret. Líq. (%)	Ret. An. (%)	Máx. Drawdown	# trades (p/dia)	# vencedores	# perdedores
Diário - Open	9.439.740,29	-0,0560	-0,1328	1.473.343,57	144 (0,51)	73 (50,7%)	71 (49,3%)
Intraday 1h	6.132.095,39	-0,3868	-0,0784	4.437.816,99	344 (1,22)	163 (47,39%)	181 (52,61%)
Intraday 15min	1.981.743,76	-0,8018	-0,7688	8.715.910,33	866 (3,64)	253 (29,22%)	613 (70,78%)

	Estratégia 2 - Close, High, Low						
	Valor Final Carteira	Ret. Líq. (%)	Ret. An. (%)	Máx. Drawdown	# trades (p/dia)	# vencedores	# perdedores
Diário - Open	12.281.378,72	0,2281	0,2475	4.370.400,42	6 (0,021)	5 (83,33%)	1 (16,67%)
Intraday 1h	10.274.920,10	0,0275	0,0016	1.089.428,38	8 (0,032)	5 (62,5%)	3 (37,5%)
Intraday 15min	10.119.944,84	0,0120	0,0013	2.027.920,44	14 (0,06)	10 (71,43%)	4 (28,57%)

	Estratégia 3 - Close, High, Low, Order						
	Valor Final Carteira	Ret. Líq. (%)	Ret. An. (%)	Máx. Drawdown	# trades (p/dia)	# vencedores	# perdedores
Diário - Open	9.451.628,88	-0,0548	-0,1289	1.899.086,92	130 (0,46)	53 (40,77%)	77 (59,23%)
Intraday 1h	6.991.828,04	-0,3008	-0,0634	4.386.450,64	247 (0,87)	107 (43,32%)	140 (56,68%)
Intraday 15min	3.670.107,18	-0,6330	-0,0308	7.182.006,23	468 (1,65)	131 (27,9915%)	337 (72,0085%)

Tabela 5.4: Tabela de resultados da aplicação do *Trading System*, com comparação entre as estratégias

ID	Tipo	Data/Hora	Preço Unit.	Quantidade	Valor Total (R\$)	Obs.
#1	Buy	08/08/2011 10:00:00	50.635,98	100	5.063.607,54	
	Reset	09/08/2011 10:00:00	50.426,22	100	5.042.632,33	Trade encerrado = buy
#2	Buy	18/08/2011 10:00:00	52.365,31	100	5.236.541,30	
	Reset	01/09/2011 10:00:00	58.278,46	100	5.827.856,32	Trade encerrado = buy
#3	Sell	01/09/2011 10:00:00	58.278,46	100	5.827.856,32	
	Reset	22/09/2011 10:00:00	54.246,28	100	5.424.638,21	Trade encerrado = sell
#4	Buy	22/09/2011 10:00:00	54.246,28	200	10.849.266,42	
	Reset	27/10/2011 10:00:00	59.067,97	200	11.813.604,58	Trade encerrado = buy
#5	Sell	27/10/2011 10:00:00	59.067,97	100	5.906.807,29	
	Reset	01/11/2011 10:00:00	56.292,96	100	5.629.305,67	Trade encerrado = sell
#6	Buy	01/11/2011 10:00:00	56.292,96	200	11.258.601,34	
	Reset	15/08/2012 10:00:00	58.300,14	200	11.660.038,30	Trade encerrado = buy

Tabela 5.5: Tabela com os trades realizados pela melhor estratégia, no período analisado

algumas situações e pelos resultados não tão favoráveis em termos de acerto do *order*.

Em termos de gráficos, são inseridos aqui os gráficos de *backtesting* na figura 5.11, onde são exibidos os pontos de compra e venda durante o período analisado, com auxílio da tabela 5.5 (referente aos trades realizados). Por fim, é exibido, na figura 5.12, a comparação da evolução dos preços (em azul) e da evolução do Patrimônio Líquido (PL) da carteira (em verde).

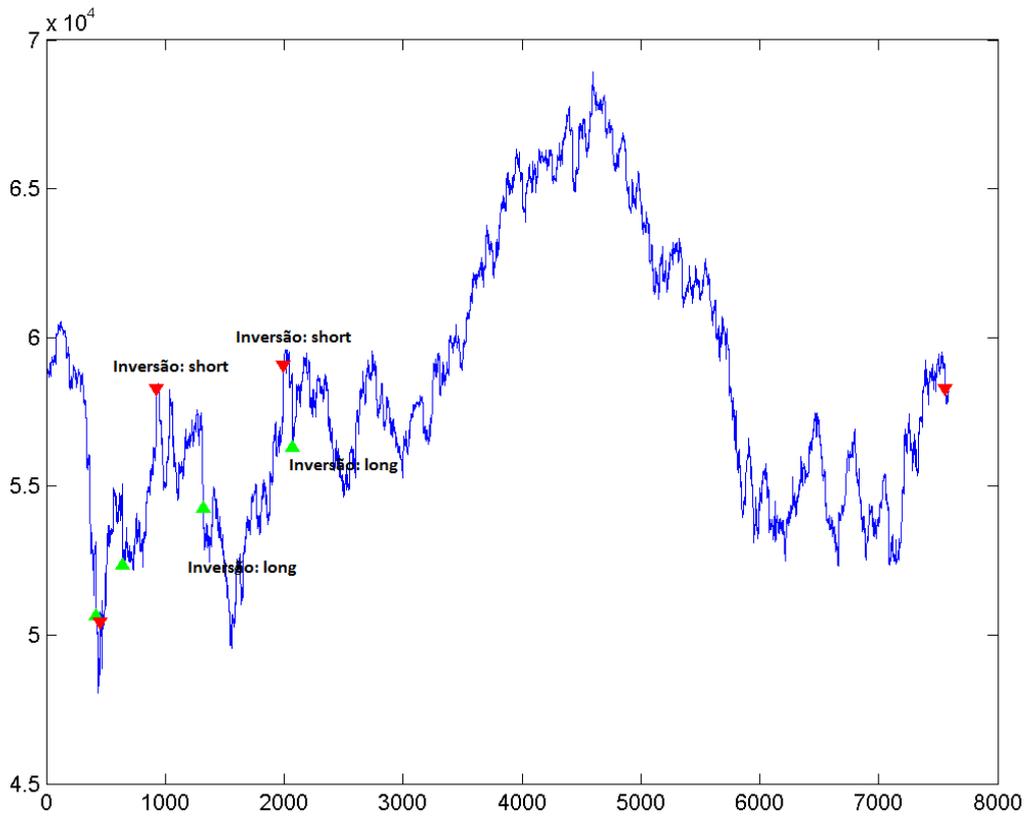


Figura 5.11: Gráfico de *backtesting*, com os pontos de compra (triângulos verdes) e de venda (triângulos vermelhos) da melhor estratégia

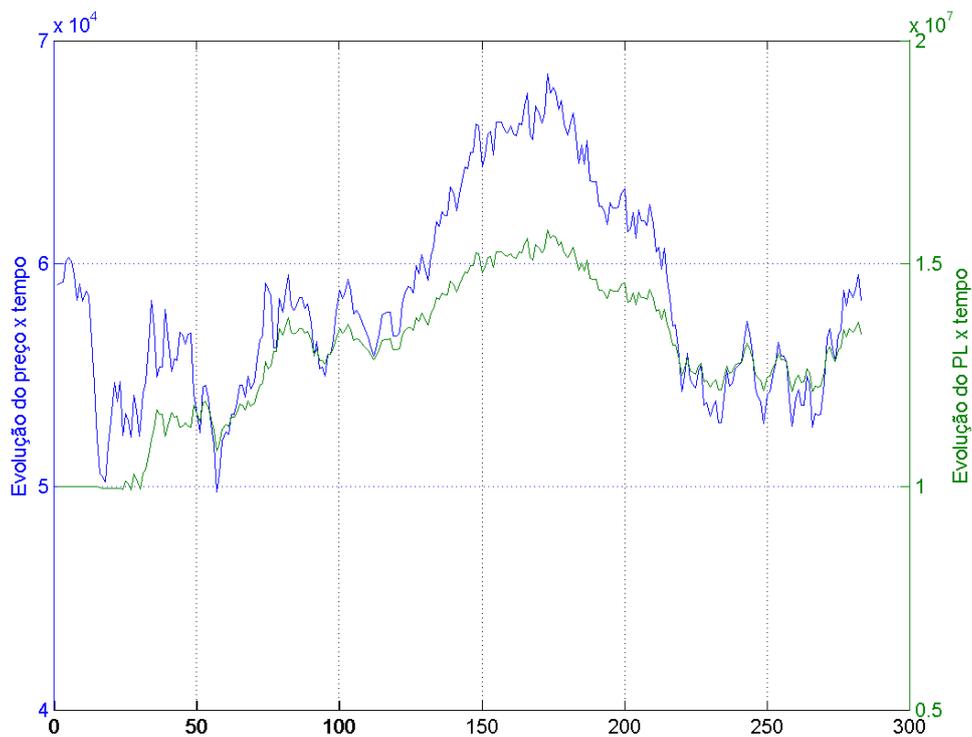


Figura 5.12: Gráfico comparativo da evolução dos Preços e do PL da carteira

5.2.4 Tempo de execução total

Em termos de estatísticas de tempo de execução, o tempo total do experimento geral foi de 25,1 horas (ou 1506 minutos), dada a execução de todos os 29 blocos gerados pelo janelamento. Em cada bloco, são 4 saídas, que pedem a execução de 27 RNAs, devido à variação de parâmetros comentada no capítulo 4, de três valores possíveis para três diferentes parâmetros.

Isso significa que temos 27 RNAs x 4 saídas x 29 blocos de janela, resultando num total de 3132 RNAs. Subtraindo do tempo total de execução o valor de 29min, referente à execução do pré-processamento de todas as séries (aproximadamente 1 minuto por bloco) e considerando o tempo total de treinamento, 25,1 horas, é possível afirmar que cada RNA levou 2,12 minutos para executar, em média. Para a execução de um único bloco, temos 108 RNAs a rodar, com um total estimado de processamento de aproximadamente 229min.

Sugerindo um cenário hipotético (não implementado nesta pesquisa), se for alcançado um bom índice de paralelização do processo, em que são executadas RNAs em 18 processos (*threads*) diferentes, teremos 6 RNAs executadas em cada processo, o que daria aproximadamente 12,8 min por processador. Adicionando 1min para o processamento e 1min para agrupamento das respostas e finalização do processamento das RNAs, teríamos menos de 15min de execução total do bloco. Sendo este o valor mínimo de janela de avaliação do *Trading System* adotado aqui, significa que poderíamos buscar realizar o treinamento na janela mínima de tempo considerada aqui (15min), o que poderia ser útil em algum sistema de mais alta frequência (HFT).

De forma óbvia, este estudo faz sentido apenas se o recebimento dos dados necessários para treinamento da RNA for praticamente instantâneo. Mas, de qualquer forma, é um bom argumento para a sugestão de paralelização do processo de treinamento de RNAs como trabalho futuro.

5.3 Resultados específicos dos blocos

Esta seção foi inserida com o intuito de apresentar exemplos de resultados específicos dos blocos executados, dado que a inclusão de todos os resultados dos 29 blocos executados seria inviável e tornaria a leitura exaustiva.

Para tal, foi escolhida a série *close* do Ibovespa, selecionada no último bloco, nº 29, para apresentar os gráficos e alguns valores obtidos, de forma a compreender melhor o processo executado dentro de cada bloco. A figura 5.13, por exemplo, exhibe os plots das etapas de transformação da série, exibindo a série ainda original, a série transformada para série de retorno e a série de retorno normalizada e, para

cada item, o histograma referente.

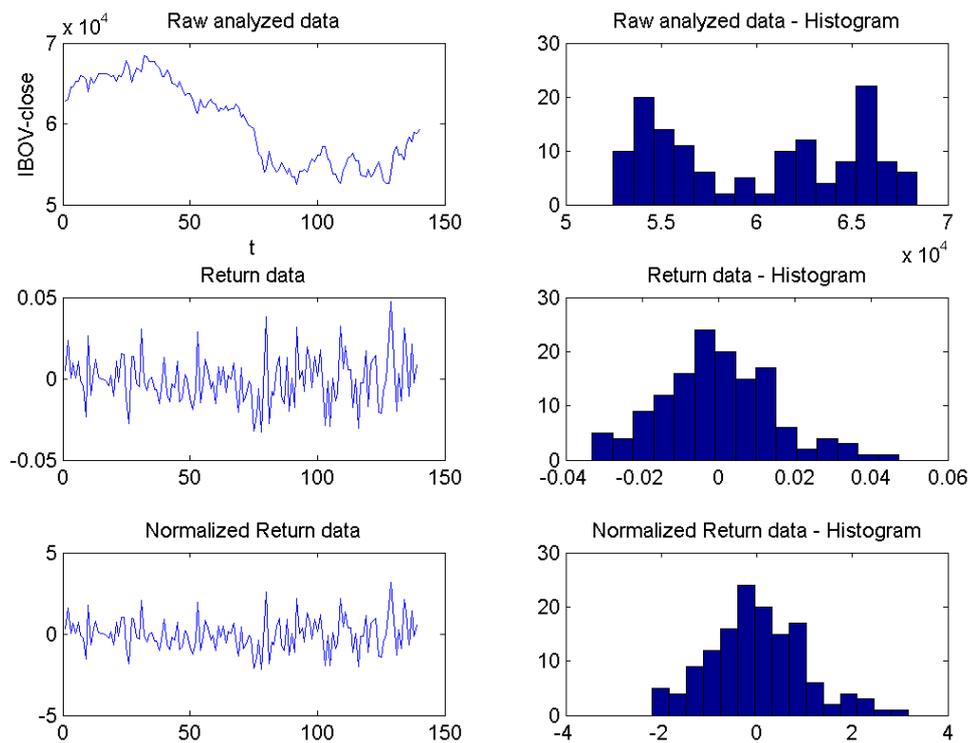


Figura 5.13: Exemplo de transformação em série de retorno e normalização, com histogramas - Ibovespa *close*

Os dois gráficos exibidos a seguir, figuras 5.14 e 5.15, exibem a forma de tratamento nos casos de extração de tendência e extração de ciclos senoidais, respectivamente. Na figura 5.14, a parte superior diz respeito à série, a linha vermelha é a reta de tendência encontrada, dada pela regressão linear $y = 0,00082193 * x - 0,063517$ e, na parte inferior, é exibida a série diminuída da reta de regressão. Por sua vez, na figura 5.15, a imagem da esquerda é o espectrograma e a da direita, a autocorrelação da série.

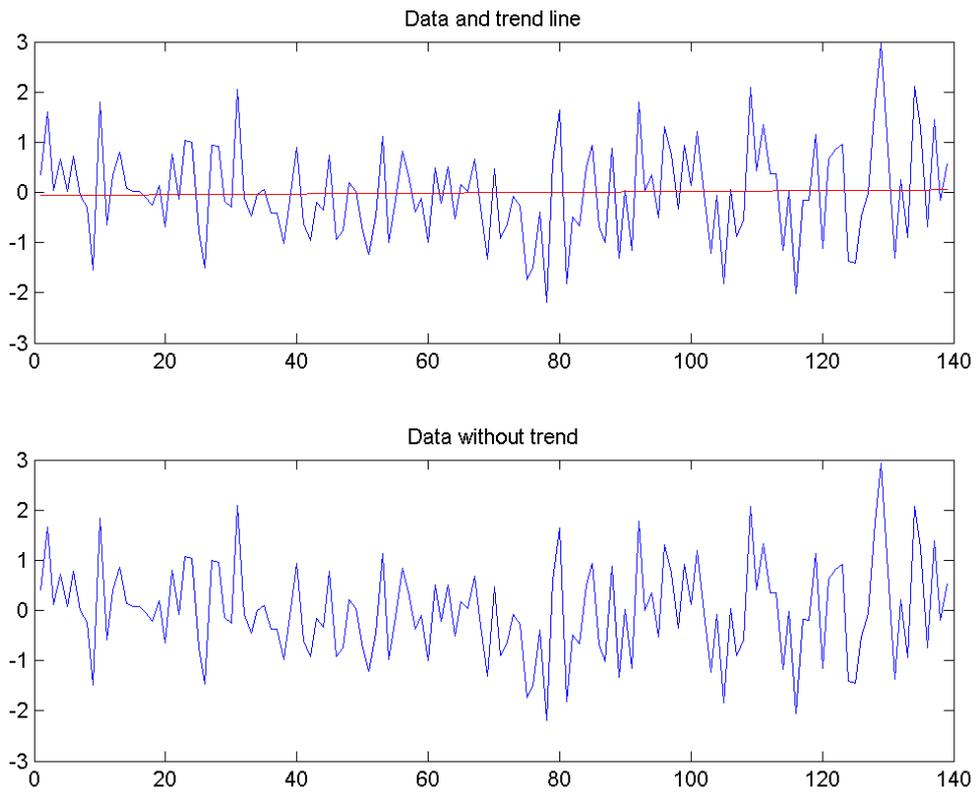


Figura 5.14: Exemplo de Extração de Tendência - Ibovespa *close*

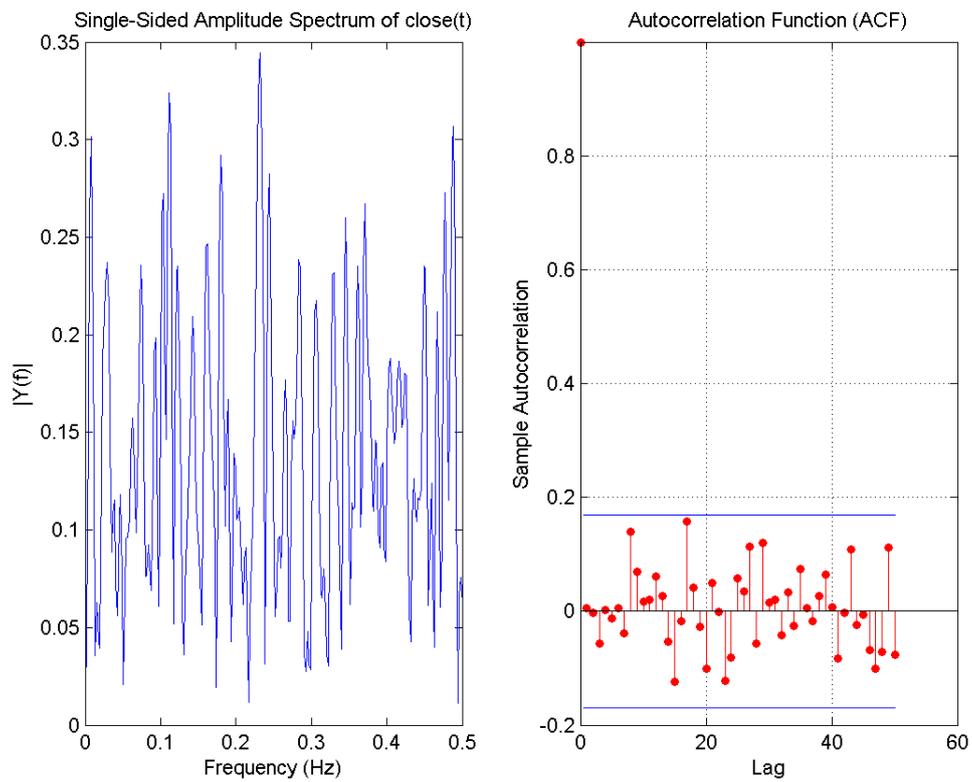


Figura 5.15: Exemplo de Espectrograma e Autocorrelação - Ibovespa *close*

O resultado desse passo não teve grande impacto no resultado final do sistema, dado que as séries já haviam sido transformadas em séries de retorno. No caso da extração de tendência, foram encontrados fatores muito próximos de 0 para o coeficiente angular e, no caso da extração de ciclos senoidais, não foram encontradas frequências significativas a remover das séries.

Outro resultado interessante é o da atenuação de intrusos (valores extremos), efetuada, da mesma forma, sobre todas as séries envolvidas no pré-processamento. O exemplo exibido a seguir é um gráfico que, na parte superior a série inicial, com os pontos extremos marcados. Nesse caso, são 10 valores (7,5%) a serem reduzidos ao seu limite superior/inferior, ou seja, $\sigma \pm 2$. A parte inferior exibe a série com os valores atenuados.

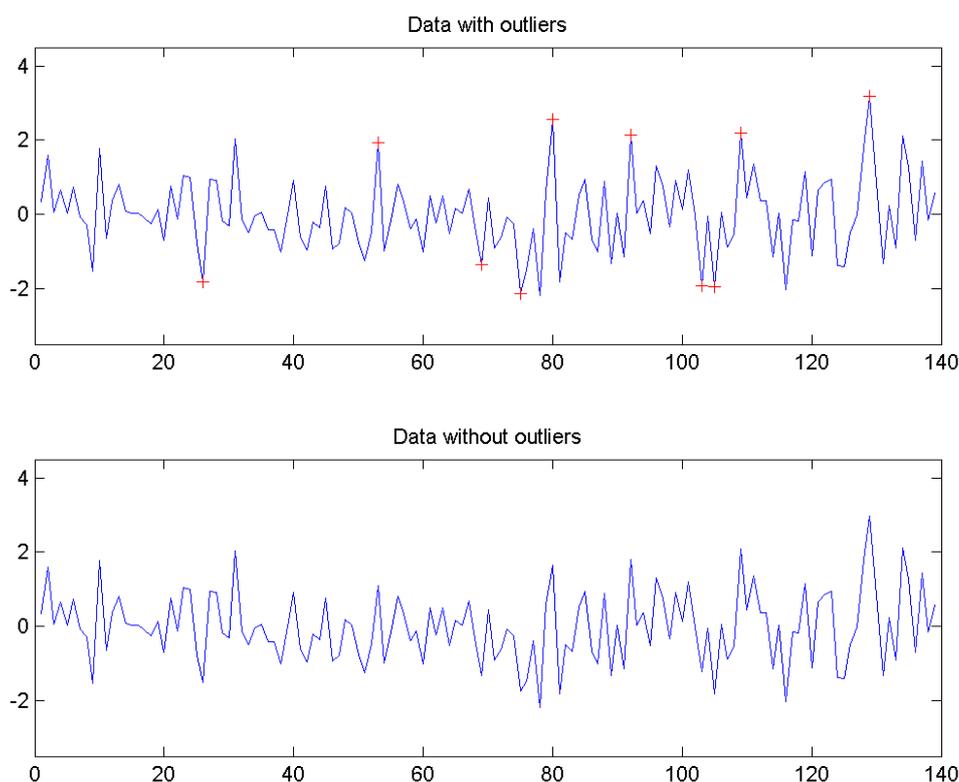


Figura 5.16: Exemplo de atenuação de valores extremos - Ibovespa *close*

Ao término da decomposição de cada série, é extraída uma tabela com as estatísticas finais sobre as mesmas. Um exemplo de tabela é exibido a seguir, na tabela 5.6, com as estatísticas obtidas nas séries originais, após a transformação em série de retorno e após a normalização, para as variáveis *close*, *high*, *low* e *order*, as saídas alvo.

Por fim, é realizada a seleção das variáveis e atrasos significativos, por meio da correlação cruzada entre as séries alvo e as séries candidatas, deslocadas para trás. A figura 5.17 a seguir exibe um exemplo de correlação cruzada entre a série *close*

		Média	Mediana	Desv. Padrão	Valor Máx.	Valor Mín.	Obliquidade	Curtose
Close	Original	60.099,92	61.020,65	5.056,52	68.394,30	52.481,40	0,0184	1,4484
	Série Ret.	-0,0003	-0,0010	0,0150	0,0472	-0,0331	0,3192	3,3091
	Normalizado	0,0000	-0,0444	1,0000	3,1782	-2,1927	0,3192	3,3091
High	Original	60.693,56	61.599,60	4.983,48	68.970,00	53.232,60	0,0033	1,4510
	Série Ret.	-0,0003	0,0000	0,0123	0,0539	-0,0271	0,6666	5,1558
	Normalizado	0,0000	0,0278	1,0000	4,4199	-2,1796	0,6666	5,1558
Low	Original	59.497,53	60.414,60	5.126,09	67.797,30	52.212,90	0,0249	1,4262
	Série Ret.	-0,0004	0,0001	0,0125	0,0382	-0,0356	-0,2619	3,5461
	Normalizado	0,0000	0,0421	1,0000	3,0895	-2,8131	-0,2619	3,5461
Order	Original	-0,1143	-1,0000	0,9970	1,0000	-1,0000	0,2301	1,0529
	Série Ret.	-0,1079	-1,0000	0,9978	1,0000	-1,0000	0,2171	1,0471
	Normalizado	0,0000	-0,8941	1,0000	1,1104	-0,8941	0,2171	1,0471

Tabela 5.6: Exemplo de tabela com as estatísticas, variáveis *close*, *high*, *low* e *order* do Ibovespa

do Ibovespa e Canadá S&P/TSX Index e pode-se perceber que foram escolhidos os *lags* de 1 e 5, neste caso, significando que Canadá S&P/TSX Index em $D - 1$ e $D - 5$ farão parte do modelo. As tabelas 5.7 e 5.8, logo a seguir, ilustram as variáveis significativas selecionadas pelo método para o bloco nº 29.

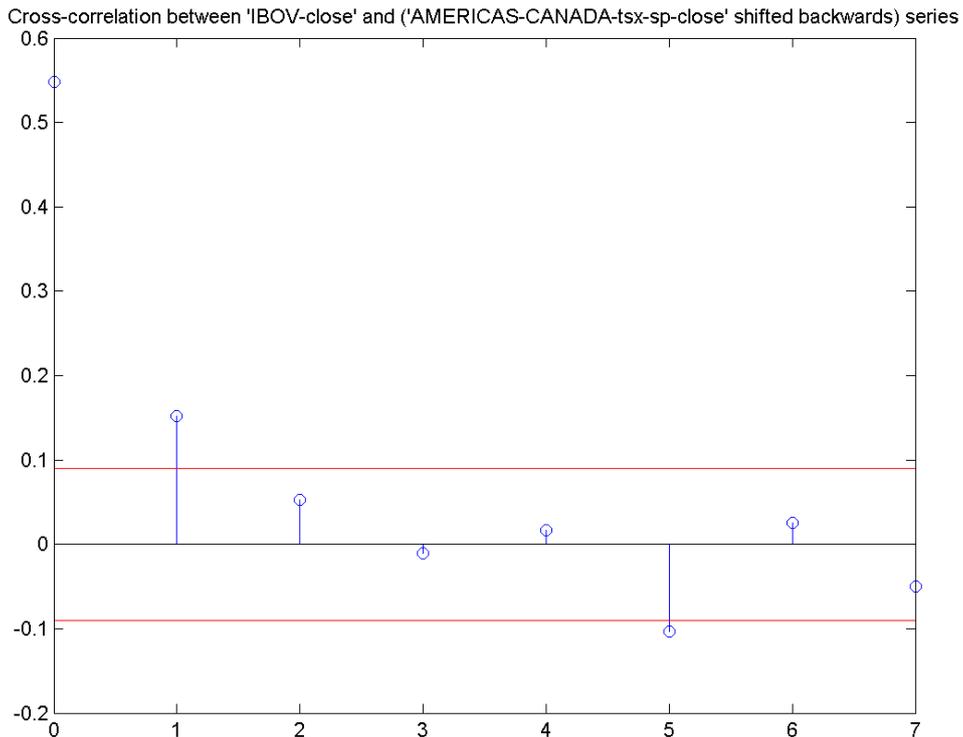


Figura 5.17: Exemplo de Correlação significativa - Canada S&P/TSX Index em relação ao Ibovespa *close* - Lags escolhidos: 1 e 5

Em termos de execução das RNAs, alguns resultados específicos são também interessantes e valem citação nesta seção. O primeiro é o acompanhamento do

Bloco	Inputs - close (25)	Inputs - high (28)
29	AMERICAS_ARGENTINA_merval_close D-1	AMERICAS_ARGENTINA_merval_close D-1
	AMERICAS_CANADA_tsx_sp_close D-1	AMERICAS_CANADA_tsx_sp_close D-1
	AMERICAS_CANADA_tsx_sp_close D-5	AMERICAS_CHILE_ipsa40_close D-1
	ASIA_CHINA_shangaiSE_SSECO_close D-0	AMERICAS_MEXICO_ipc_close D-1
	ASIA_HONGKONG_hangse_close D-2	ASIA_CHINA_shangaiSE_SSECO_close D-0
	ASIA_HONGKONG_hangse_close D-0	ASIA_HONGKONG_hangse_close D-0
	ASIA_JAPAN_nikkei_close D-1	ASIA_JAPAN_nikkei_close D-0
	ASIA_JAPAN_nikkei_close D-2	ASIA_KOREA_kospi_close D-0
	ASIA_JAPAN_nikkei_close D-0	ASIA_RUSSIA_rtsi_close D-1
	ASIA_KOREA_kospi_close D-2	ASIA_RUSSIA_rtsi_close D-0
	ASIA_KOREA_kospi_close D-0	BRA_fgv100_close D-1
	ASIA_RUSSIA_rtsi_close D-2	BRA_ibov_fut_INDV12_close D-1
	ASIA_RUSSIA_rtsi_close D-0	BRA_idiv11_close D-1
	BRA_cdi252d_close D-6	BRA_imob_close D-1
	BRA_dolar_ptax_close D-5	BRA_sml11_close D-1
	BRA_eurbri_currency_close D-4	EUR_FRANCE_CAC40_close D-1
	BRA_eurbri_currency_close D-5	EUR_GERMANY_closeax_close D-1
	EUR_FRANCE_CAC40_close D-1	EUR_ITALY_ftseMIB_close D-1
	IBOV_AT_macd_10_30_7d D-2	EUR_UK_ftse100_close D-1
	IBOV_AT_obv_100d D-7	IBOV_AT_ifr9d D-1
	IBOV_holiday_eve_br D-5	IBOV_AT_macd_10_30_7d D-3
	IBOV_holiday_usa D-7	IBOV_close D-1
	IBOV_var_daily D-4	IBOV_high D-1
	IBOV_var_max D-4	IBOV_low D-1
	USA_dowjones_close D-1	IBOV_open D-0
	IBOV_order D-1	
	USA_dowjones_industrialaverage30_close D-1	
	USA_sp500_close D-1	

Tabela 5.7: Variáveis e lags selecionados no bloco nº 29 para as RNAs das séries alvo: *Close* e *High*

Bloco	Inputs - low (29)	Inputs - order (19)
29	AMERICAS_ARGENTINA_merval_close D-1	AMERICAS_CANADA_tsx_sp_close D-1
	AMERICAS_CANADA_tsx_sp_close D-1	ASIA_CHINA_shangaiSE_SSECO_close D-0
	AMERICAS_MEXICO_ipc_close D-1	ASIA_HONGKONG_hangse_close D-7
	ASIA_CHINA_shangaiSE_SSECO_close D-0	ASIA_HONGKONG_hangse_close D-0
	ASIA_HONGKONG_hangse_close D-0	ASIA_JAPAN_nikkei_close D-7
	ASIA_JAPAN_nikkei_close D-0	ASIA_KOREA_kospi_close D-0
	ASIA_KOREA_kospi_close D-0	ASIA_RUSSIA_rtsi_close D-0
	ASIA_RUSSIA_rtsi_close D-1	BRA_cdi252d_close D-7
	ASIA_RUSSIA_rtsi_close D-0	BRA_eurbrl_currency_close D-4
	BRA_fgv100_close D-1	BRA_eurbrl_currency_close D-5
	BRA_ibov_fut_INDV12_close D-1	BRA_idiv11_close D-4
	BRA_idiv11_close D-1	EUR_FRANCE_CAC40_close D-1
	BRA_imob_close D-1	IBOV_AT_ifr9d D-4
	BRA_sml11_close D-1	IBOV_AT_macd_10_30_7d D-3
	EUR_FRANCE_CAC40_close D-1	IBOV_AT_obv_100d D-7
	EUR_GERMANY_closeax_close D-1	IBOV_holiday_eve_br D-5
	EUR_ITALY_ftseMIB_close D-1	IBOV_open D-7
	EUR_UK_ftse100_close D-1	IBOV_order D-4
	IBOV_AT_ifr9d D-1	IBOV_var_daily D-7
	IBOV_AT_macd_10_30_7d D-3	
	IBOV_AT_macd_10_30_7d D-4	
	IBOV_close D-1	
	IBOV_high D-1	
	IBOV_low D-1	
	IBOV_open D-0	
	IBOV_order D-1	
	IBOV_var_max D-1	
	USA_dowjones_industrialaverage30_close D-1	
	USA_sp500_close D-1	

Tabela 5.8: Variáveis e *lags* selecionados no bloco nº 29 para as RNAs das séries alvo: *Low* e *Order*

treinamento, que é realizado por um plot com a evolução do erro nos 3 conjuntos de divisão das RNAs: treinamento, validação e teste. A figura 5.18 pode ser conferida a seguir, sendo que o eixo x , no caso, representa a época na qual os erros são obtidos.

Por fim, na figura 5.19, logo a seguir, são comparados os resultados obtidos (reais *versus* estimados) para este bloco específico de treinamento. No caso, estes dois exemplos são referentes ao treinamento da RNA da saída Ibovespa *high*.

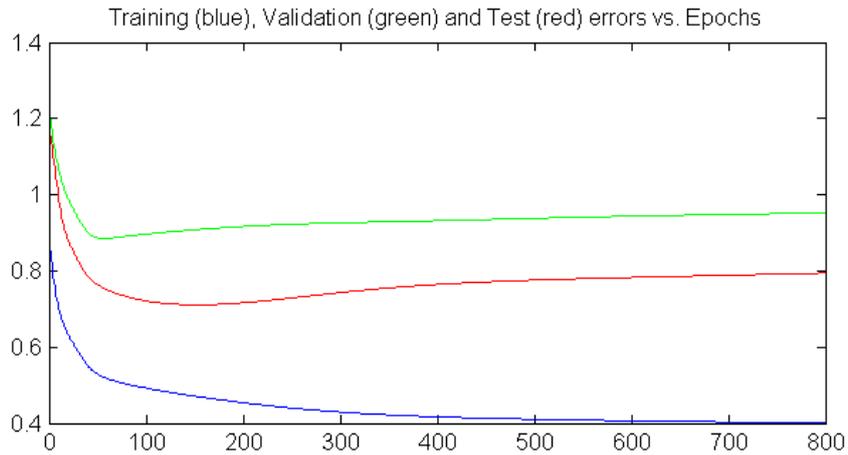


Figura 5.18: Exemplo de gráfico de acompanhamento dos erros de treinamento, validação e teste *versus* épocas de treinamento - Ibovespa *high*

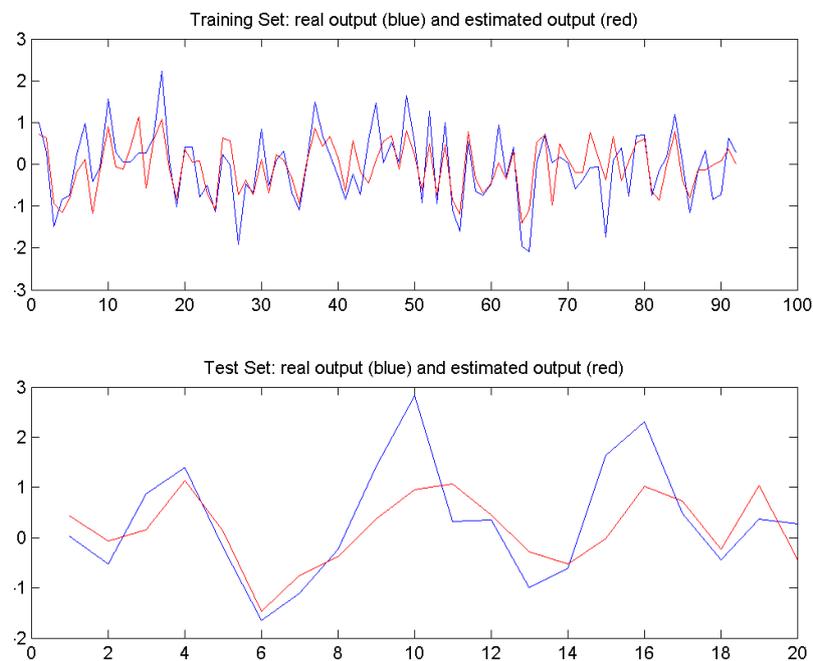


Figura 5.19: Exemplo de gráfico de comparação entre valores reais (linha azul) e estimados (linha vermelha), para os conjuntos de treinamento e teste - Ibovespa *high*

Capítulo 6

Conclusões

6.1 Conclusão

O objetivo principal deste trabalho, conforme explicado no capítulo 1, foi "avaliar o desempenho de uma estratégia automatizada de negociação, baseada em Redes Neurais Artificiais, que realiza predições sobre as principais séries temporais do índice Ibovespa".

Com o intuito de atingir o objetivo descrito, foram estudadas algumas das ferramentas necessárias para realizar operações lucrativas no mercado brasileiro, baseadas em predições obtidas por RNAs. Isto pôde ser obtido através da investigação de padrões do mercado, por intermédio do índice Ibovespa, principal referência para a bolsa de valores de São Paulo (BM&FBovespa). Foram quatro séries alvo estudadas: *close*, *high*, *low* e *order*, representando respectivamente: fechamento, máxima, mínima e a ordem em que máxima e mínima ocorrem, todos em valores diários.

Para as predições, foi escolhida a implementação de RNAs, que é considerado um método clássico de regressão (e classificação). Além de suas inerentes características, há uma extensa base literária reforçando sua escolha. Foi utilizada a mesma arquitetura de RNAs para cada um dos experimentos, com uma camada intermediária e com um esquema de variação de parâmetros, para o número de neurônios na camada intermediária, valor máximo inicial (em absoluto) para os pesos sinápticos e a taxa de aprendizado, por serem estes os parâmetros mais afetados pela alta variação das possibilidades de entradas nos modelos.

A maior diferença entre cada RNA é os dados de entrada considerados, que foram obtidos após um pré-processamento sobre diversas séries temporais de bolsas mundiais e indicadores domésticos. Os atrasos significativos de cada uma destas foram determinados após os cálculos de correlação cruzada entre as séries estudadas e as séries alvo.

Foram escolhidos quatro critérios para avaliação dos modelos: RMSE, MAPE,

Correlação e Acurácia. Os resultados obtidos para o *close* mostram valor baixo para o erro RMSE, aliado a uma acurácia de aproximadamente 60% da série real. Os resultados para *high* e *low* foram ainda melhores. Por outro lado, a série *order* se mostrou de difícil predição, obtendo valores altos de falsos negativos e falsos positivos. Tais resultados reforçam a ideia de que é possível realizar uma boa predição sobre o comportamento do mercado para $D + 1$, inclusive podendo superar *benchmarks* importantes, como o próprio índice Ibovespa.

Antes de inserir os valores na RNA, as séries sofreram um processo de decomposição, com o objetivo de melhorar o processo de treinamento das RNAs e diminuir a possibilidade de especialização em padrões indesejados das redes. Todo esse pré-processamento e, por consequência, a aplicação das RNAs, foram sujeitos a um esquema de janelamento, criado a partir de uma especialização do *walk-forward training*, de forma a aproximar a aplicação da realidade.

Com os resultados dos experimentos, foi possível implementar e utilizar um sistema automatizado para negociação no mercado (*Trading System*). A aplicação das RNAs fica mais bem avaliada dessa forma, não somente obtendo os resultados clássicos de erros de predição, como também os resultados financeiros da aplicação das predições em estratégias. A intenção é determinar a possibilidade deste se tornar um método lucrativo para investidores e gestores de fundos de investimento, a partir de uma aplicação hipotética com um ativo real operado no mercado, no caso, contratos de Futuro de Ibovespa.

Foram criadas três estratégias como meio de simular as aplicações, onde uma delas se saiu melhor do que as outras, provavelmente por ter efetuado menor quantidade de trades e, os executados, foram mais acertados do que os de outras estratégias. O sistema considera também restrições e taxas do mercado financeiro real, como *slippage*, corretagem, Imposto de Renda, gerenciamento de risco (por meio de *stop-loss*) e condições de negociação.

A aplicação das estratégias levou em conta as predições diárias para aplicação em estratégias também diárias ou *intraday*, ou seja, levando em conta as cotações durante o dia para a tomada de decisão de compra ou venda. Os resultados obtidos pelo *Trading System* foram interessantes porque algumas estratégias se mostraram lucrativas, mesmo considerando que, no período estudado, o preço do ativo apresentou queda de -17,05%.

Resumindo, podem ser destacados como diferenciais desta pesquisa:

- Utilização de um modelo adaptativo que se renova, por meio do janelamento, a cada período de aplicação;
- Avaliação da aplicação de correlação parcial para obtenção das variáveis significativas;

- Uso de diferentes abordagens de RNAs para problemas de regressão e classificação;
- Simulação da operação real e fiel do mercado por meio de um Trading System e um ativo real operado no mercado (Futuro de Ibovespa).

6.2 Trabalhos Futuros

Alguns pontos trabalhados aqui merecem uma investigação mais detalhada, como, por exemplo, a parte de estratégias de operação. Pode-se pensar em incluir variáveis de Análise Técnica nas estratégias como forma de incrementar ou limitar a decisão inicial, tomada com base nos valores de previsões de cada RNA.

Outro ponto que merece uma atenção maior é em relação ao desempenho das RNAs. É sugerida a investigação de utilização de outras formas de variação de parâmetros, considerando uma taxa de aprendizado adaptativa e, possivelmente, descartando a necessidade de variação deste parâmetro. Sugere-se também a tentativa de implementação de novas arquiteturas de RNA ou, até mesmo, outros métodos de *machine learning*, como SVR, SVM, sistemas híbridos, modelos neuro-fuzzy e RNAs com algoritmo genético, de forma a melhorar os resultados de previsão. Outra possibilidade de aperfeiçoamento pode também ser descoberta na utilização de séries temporais financeiras não consideradas neste trabalho, como outras variáveis de Análise Técnica ou outros índices mundiais.

Por fim, sugere-se aplicar na prática o estudo proposto no capítulo 5, sobre a paralelização do processo. De acordo com as estimativas de cálculo, pode ser possível, dado um número determinado de processos executando em paralelo, treinar um bloco inteiro em menos de 15min, que é o valor mínimo de janela *intraday* utilizado aqui. Pode ser um possível caminho para a aplicação de um sistema de previsões *intraday*, com atualização a cada deslocamento de tempo.

Referências Bibliográficas

- [1] “The World’s Biggest Stock Exchanges”. <http://www.forbes.com/pictures/eddk45iglh/bmf-bovespa-brazil/>. Accessed: 2014-03-13.
- [2] . “BM&FBOVESPA - Seu futuro bem investido”. <http://www.bmfbovespa.com.br/home.aspx?idioma=pt-br>, . Accessed: 2014-03-13.
- [3] . “O que é Hedge?” <http://www.itauinvestnet.com.br/itauinvestnet/fundos/entenda/conteudo.aspx?op=35>, . Accessed: 2014-03-13.
- [4] ABU-MOSTAFA, Y. S., ATIYA, A. F., 1996, “Introduction to Financial Forecasting”, *Applied Intelligence*, v. 6, pp. 205–213.
- [5] ALBERT Y. ZOMAYA, JAMES A. ANDERSON, D. B. F. G. J. M., ROZENBERG, G., 2001, “Nonconventional computing paradigms in the new millennium: A roundtable”, *Computing in Science and Engg*, pp. 82–99.
- [6] ATSALAKIS, G. S., VALAVANIS, K. P., 2009, “Surveying stock market forecasting techniques Part II Soft computing methods”, *Expert Systems with Applications*, v. 36, n. 3, Part 2, pp. 5932 – 5941. ISSN: 0957-4174. doi: <http://dx.doi.org/10.1016/j.eswa.2008.07.006>.
- [7] BENGIO, Y., LECUN, Y., 2007, “Scaling Learning Algorithms towards AI”. In: Bottou, L., Chapelle, O., Decoste, D., et al. (Eds.), *Large-Scale Kernel Machines*, MIT Press.
- [8] BM&FBOVESPA, 2010, “Introdução Ao Mercado De Capitais”, .
- [9] BOX, G. E. P., JENKINS, G. M., 1994, *Time Series Analysis: Forecasting and Control*. 3rd ed. Upper Saddle River, NJ, USA, Prentice Hall PTR. ISBN: 0130607746.
- [10] BRAGA, P., 2006, *Previsão do IBOVESPA Utilizando Modelos Híbridos*. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

- [11] BRUNO C. BARROSO, GUSTAVO P. HANAOKA, F. D. P. R. T. N. C., “Uma ferramenta computacional eficiente para identificação de melhores gatilhos de compra e venda de ativos financeiros”, .
- [12] CALÔBA, L. P., 2002, “Introdução ao uso de Redes Neurais na Modelagem de Sistemas Dinâmicos e Séries Temporais”, *Livro de Mini Cursos do XIV Congresso Brasileiro de Automática*.
- [13] CHANDE, T. S., 2001, *Beyond Technical Analysis: How to Develop and Implement a Winning Trading System*. Wiley. ISBN: 978-0-471-41567-1.
- [14] CHEN, Y., M. S. H. K. . H. J., 2007, “Genetic network programming with Sarsa Learning and its application to creating stock trading rules”, *Proceedings of the 2007 IEEE congress on evolutionary computation*, pp. 220–227.
- [15] CHEN, Y., M. S. K. S. H. K., 2009, “A genetic network programming with learning approach for enhanced stock trading model”, *Expert Systems with Applications*, pp. 12537–12546.
- [16] CHENOWETH, T., OBRADOVIC, Z., 1996, “A multi-component nonlinear prediction system for the S&P 500 index”, *Neurocomputing*, v. 10, n. 3, pp. 275 – 290. ISSN: 0925-2312. doi: [http://dx.doi.org/10.1016/0925-2312\(95\)00109-3](http://dx.doi.org/10.1016/0925-2312(95)00109-3). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0925231295001093>>. Financial Applications, Part {II}.
- [17] CORTEZ, P. A. R., 1997, *Algoritmos Genéticos e Redes Neurais na Previsão de Séries Temporais*. Dissertação de M.Sc., Universidade do Minho, Braga, Portugal.
- [18] COSTA, V., 2008. “Análise De Pré-Processamento No Desempenho De Previsão De Séries Temporais Financeiras”. .
- [19] CYBENKO, G., 1989, “Approximation by superpositions of a sigmoidal function”, *Mathematics of Control, Signals and Systems*, v. 2, n. 4, pp. 303–314. ISSN: 0932-4194. doi: 10.1007/BF02551274. Disponível em: <<http://dx.doi.org/10.1007/BF02551274>>.
- [20] DA S. SOARES, A., 2008, *Predição de Séries Temporais Econômicas por meio de Redes Neurais Artificiais e Transformada Wavelet: Combinando Modelo Técnico e Fundamentalista*. Tese de Mestrado, USP.
- [21] DANTAS, A. C. H., 2007, *Sistema de Monitoração de Qualidade de Dados*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

- [22] EHLERS, R. S., 2009, *Análise de Séries Temporais*. Relatório técnico.
- [23] FARNSWORTH, G. V., K. J. A. O. A. S. P. R. J., 2004, *Successful technical trading agents using genetic programming*. Tech. Rep. SAND2004-4774.
- [24] G. ZHANG, B. E. PATUWO, M. Y. H., 1998, “Forecasting with artificial neural networks: The state of the art”, *International Journal of Forecasting*, v. 14, pp. 35–62.
- [25] GARCÍA-FERRER, A., 2012, “On Granger’s predictability of financial markets in theory and practice”, *International Journal of Forecasting*, v. 28, n. 1, pp. 121 – 127. ISSN: 0169-2070. doi: <http://dx.doi.org/10.1016/j.ijforecast.2011.02.011>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0169207011000379>. Special Section 1: The Predictability of Financial Markets Special Section 2: Credit Risk Modelling and Forecasting.
- [26] GOMIDE, P., MILIDIÚ, R. L., 2010, “Assessing Stock Market Time Series Predictors Quality through a Pairs Trading System”. In: *SBRN*, pp. 133–139.
- [27] GRANGER, C. W. J., 1992, “Forecasting stock market prices: Lessons for forecasters”, *International Journal of Forecasting*, v. 8, n. 1, pp. 3 – 13. ISSN: 0169-2070. doi: [http://dx.doi.org/10.1016/0169-2070\(92\)90003-R](http://dx.doi.org/10.1016/0169-2070(92)90003-R). Disponível em: <http://www.sciencedirect.com/science/article/pii/S016920709290003R>.
- [28] HAYKIN, S., 2000, *Redes Neurais Artificiais: Princípios e Práticas*. São Paulo, Bookman.
- [29] HELSTROM, T., HOLMSTROM, K., 1998, “Predicting the Stock Market”, *Opuscula ISRN HEV-BIB-OP-26-SE*.
- [30] KAASTRA, I., BOYD, M., 1996, “Designing a Neural Network for Forecasting Financial and Economic Time Series”, *Neurocomputing*.
- [31] KALYVAS, E., 2001, *Using Neural Networks and Genetic Algorithms to Predict Stock Market Returns*. Tese de Mestrado, University of Manchester.
- [32] KOHAVI, R., PROVOST, F., 1998, “Glossary of Terms”, *Machine Learning*, v. 30, n. 2-3 (February), pp. 271–274. Disponível em: <http://www.springerlink.com/content/x105525142v51t20/>.

- [33] KROLLNER, B., V. B., FINNIE, G., 2010, “Financial time series forecasting with machine learning techniques: A survey”, *European Symposium on Artificial Neural Networks: Computational and Machine Learning*.
- [34] L. C. MARTINEZ, D. N. DA HORA, J. R. D. M. P. W. M. J., PAPPA, G. L., 2009, “From an Artificial Neural Network to a Stock Market Day-trading System: A Case Study on the BM&F BOVESPA”. In: *Proceedings of the 2009 International Joint Conference on Neural Networks, IJCNN’09*, pp. 3251–3258, Piscataway, NJ, USA. IEEE Press. ISBN: 978-1-4244-3549-4. Disponível em: <<http://dl.acm.org/citation.cfm?id=1704555.1704743>>.
- [35] L. S. MACIEL, R. B., 2010, “Neural Networks Applied To Stock Market Forecasting: An Empirical Analysis”, *Journal of the Brazilian Neural Network Society*, v. 8, n. 1, pp. 3 – 22.
- [36] LAWRENCE, R., 1997, “Using Neural Networks to Forecast Stock Market Prices”, .
- [37] MADHAVARAM, G. R., 2013, *Statistical Arbitrage Using Pairs Trading With Support Vector Machine Learning*. Dissertação de M.Sc., Saint Mary’s University, Halifax, Nova Scotia, Canada.
- [38] MALKIEL, B., 1999, *A random walk down wall street*. W. W. Norton & Company.
- [39] MALKIEL, B. G., FAMA, E. F., 1970, “Efficient capital markets: A review of theory and empirical work”, *The journal of Finance*, v. 25, n. 2, pp. 383–417.
- [40] MANTRI, J. K., 2013, “Comparison between SVM and MLP in Predicting Stock Index Trends”, ISSN: 2319-6386.
- [41] MCCULLOCH, W. S., PITTS, W. A., 1943, “Logical Calculus of the Ideas Immanent in Nervous Activity”, .
- [42] MIRANDA, M. E., 2011, *Redes Neurais Aplicadas à Estimação de Séries Temporais Financeiras*. Monografia, Escola Politécnica/UFRJ, Rio de Janeiro, RJ, Brasil.
- [43] MORETTIN, P. A., TOLOI, C. M. C., 2006, *Análise de Séries Temporais*. São Paulo, Edgard Blucher.
- [44] MÁSSON, E., WANG, Y.-J., 1990, “Introduction to Computation and Learning in Artificial Neural Networks”, .

- [45] MUELLER, A., 1996, *Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário*. Tese de Mestrado, UFSC.
- [46] MURARKA, M. S. A. D. P. D., 2013, “Stock Price Forecasting : Comparison of Short Term and Long Term Stock Price Forecasting using Various Techniques of Artificial Neural Networks”, *International Journal of Advanced Research in Computer Science and Software Engineering*, v. 3, n. 6. ISSN: 2277 128X.
- [47] MURPHY, J., 1999, *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance Series. New York Institute of Finance. ISBN: 9780735200661.
- [48] ORTEGA, L. F., 2013, “A Neuro-wavelet Model for the Short-term Forecasting of High-frequency Time Series of Stock Returns”, *Journal of Forecasting*.
- [49] OSÓRIO, R. H. C., 2008, *Estimação do Índice IBOVESPA de Abertura A Partir de Outros Índices Mundiais*. Monografia, Escola Politécnica/UFRJ, Rio de Janeiro, RJ, Brasil.
- [50] PALIWAL, M., KUMAR, U. A., 2009, “Neural networks and statistical techniques: A review of applications”, *Expert Systems with Applications*, v. 36, n. 1, pp. 2 – 17. ISSN: 0957-4174. doi: <http://dx.doi.org/10.1016/j.eswa.2007.10.005>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417407004952>.
- [51] RIPLEY, B. D., 1996. “Pattern Recognition via Neural Networks” . .
- [52] ROSENBLATT, F., 1958, “The perceptron: a probabilistic model for information storage and organization in the brain.” *Psychological review*, v. 65, n. 6, pp. 386.
- [53] SCHOPF, E. C., 2007, *Método Neuro-estatístico para Predição de Séries Temporais Ruidosas*. Dissertação de M.Sc., Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.
- [54] WOLPERT, D. H., 1996, “The Lack of a Priori Distinctions Between Learning Algorithms”, *Neural Comput.*, v. 8, n. 7 (out.), pp. 1341–1390. ISSN: 0899-7667. doi: 10.1162/neco.1996.8.7.1341. Disponível em: <http://dx.doi.org/10.1162/neco.1996.8.7.1341>.
- [55] YOSHIMOTO, N. “Fundos Quantitativos”. <http://www.webfinder.com.br/disclosure/HTM/D128-art3.htm>. Accessed: 2014-03-13.

- [56] ZABBAH, I., PARTOVI, E., “Enhancing an Automated Trading Strategy Using Artificial Neural Networks”, *International Conference on Computer Engineering and Network Security (ICCENS'2012)*.