

**UNIVERSIDADE DE CAXIAS DO SUL
CENTRO DE CIÊNCIAS EXATAS E DA TECNOLOGIA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

CÁSSIO CRISTIANO THIELE

**COMBINANDO SISTEMAS ESPECIALISTAS NA PREVISÃO DE SÉRIES
TEMPORAIS FINANCEIRAS**

**CAXIAS DO SUL
2015**

CÁSSIO CRISTIANO THIELE

**COMBINANDO SISTEMAS ESPECIALISTAS NA PREVISÃO DE SÉRIES
TEMPORAIS FINANCEIRAS**

Monografia apresentada como requisito para a
obtenção do grau de Bacharel em Ciência da
Computação da Universidade de Caxias do Sul.

Orientador Prof. Dr. André Gustavo Adami

CAXIAS DO SUL
2015

Aos meus pais Mário e Mirtes, à minha irmã Joice, à minha namorada Cibéle e a toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

AGRADECIMENTOS

A Deus por minha vida, família, amigos e por ter me dado saúde e força para superar as dificuldades.

À minha mãe, Mirtes Teresinha Vogel, por ter me criado com todo o amor e carinho, e por me dar apoio e incentivo nas horas difíceis, de desânimo e cansaço.

Ao meu pai, Mário Luiz Thiele, por sempre acreditar em mim e pelos ensinamentos dados ao longo da vida.

À minha irmã, Joice Tailene Thiele, pelas palavras de incentivo e pela parceria nas madrugadas de TCC. Foi difícil mas vencemos juntos.

À minha namorada, Cibéle Kolb, por sempre estar ao meu lado, por todo amor e carinho e por ter me aturado nos momentos de estresse.

Ao professor orientador, Dr. André Gustavo Adami, por dedicar seu tempo para me ajudar e me guiar no decorrer deste trabalho, me dando todo o apoio necessário.

A todos os professores da Universidade pelos inúmeros ensinamentos dados ao longo deste curso.

A todos os meus colegas de graduação, em especial aos colegas Matheus Pereira, Joel Fabiano Hansen e André Luis S. Reckziegel pelo companheirismo e troca de conhecimentos nas disciplinas que cursamos juntos.

A todos os professores do Ensino Fundamental e Médio que, sem a menor dúvida, contribuíram muito na minha formação.

Aos meus colegas de trabalho, em especial ao meu amigo e chefe Ruben Hansen, pelo apoio dado durante esta caminhada.

À toda minha família e a todas as pessoas que, direta ou indiretamente, ajudaram a tornar o curso de graduação e este trabalho possível.

“Julgue seu sucesso pelas coisas que você teve que renunciar para conseguir.”
(Dalai Lama)

RESUMO

A projeção do comportamento das ações é de fundamental importância para investidores obterem bons retornos financeiros no mercado acionário. A possibilidade de previsão de valores futuros de uma série temporal financeira está baseada em uma das principais premissas da escola de análise técnica de ações: preços futuros são uma repetição de preços passados. Modelos previsores podem ser criados através de sistemas especialistas, que são sistemas que detêm um conhecimento de qualidade especialista sobre um determinado assunto. Este trabalho teve como finalidade criar um modelo previsor, baseado na combinação de sistemas especialistas, para prever a tendência de movimentação do preço de um selecionado de ações. Redes Neurais Artificiais foram utilizadas para implementar os sistemas especialistas. Dados históricos e de indicadores de análise técnica foram utilizados como entrada para as redes neurais. A previsão de tendência, de alta ou de baixa no preço, foi para um horizonte de 1 dia, o próximo dia da série temporal. O sistema previsor final alcançou uma taxa de acerto variando entre 55% e 58% para as 3 ações analisadas. Ao realizar uma simulação de investimento, observou-se que sua utilização possibilitaria a obtenção de lucros de 13% a 24%, no período de 1 ano. Considerando que estas ações tiveram queda de 15% a 30% nesse mesmo período, pode-se dizer que o sistema previsor apresentou um bom desempenho.

Palavras-chaves: Previsão, Tendência, Mercado de Ações, Sistemas Especialistas, Redes Neurais Artificiais.

ABSTRACT

The prediction of the stock behavior is of fundamental importance for investors to obtain good financial returns in the stock market. The ability to forecast future values of financial time series is based on one of the main premises of the school of technical stock analysis: future prices are a repetition of past prices. Forecast models can be developed through expert systems, which are systems that have expert knowledge on a particular subject. This work aims to create a predictor model based on the combination of expert systems to predict the price movement trend of a selected stock. Artificial neural networks are used to implement expert systems. Historical data and analysis techniques indicators were used as input to the neural networks. The forecast trend, high or low in price, is for a horizon of one day, that is, the next day in the time series. The trend forecast system achieved a hit rate ranging between 55% and 58% for the 3 analyzed stocks. When performing an investment simulation, we found that the system would return between 13% and 24% in profits over the 1-year period. Given that the stocks used in the experiments fell by 15% to 30% during the same period, it is possible to claim that the predictor system performed well. Further research could be addressed to answer how the trend is affected by each one of the input features.

Key-words: Forecast, Trend, Stock Market, Expert Systems, Artificial Neural Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – Intersecção de Médias Móveis - ação VALE5 da empresa Vale.	30
Figura 2 – Bandas de Bollinger - ação PETR4 da empresa Petrobrás.	31
Figura 3 – Convergência e Divergência da Média Móvel (MACD) - ação PETR4 da empresa Petrobrás.	32
Figura 4 – ROC_{10} - ação PETR4 da empresa Petrobrás.	33
Figura 5 – Série Temporal da ação PETR4 da Empresa Petrobrás, de 22 de outubro de 2013 a 17 de outubro de 2014.	35
Figura 6 – Arquitetura genérica de sistemas de previsão.	36
Figura 7 – Modelo genérico de um neurônio artificial.	41
Figura 8 – Arquitetura de RNA-MLP com duas camadas ocultas.	42
Figura 9 – Gráfico da Função Sigmóide.	45
Figura 10 – Gráfico da Função Tangente Hiperbólica.	45
Figura 11 – Um exemplo de um problema de 2 classes linearmente separáveis com dois classificadores lineares possíveis.	46
Figura 12 – Arquitetura do Sistema Previsor Proposto.	54
Figura 13 – Gráfico de dispersão da relação entre taxa de acerto e taxa de retorno da simulação de investimento, obtido através dos experimentos realizados na RNA com Variáveis Dependentes.	67
Figura 14 – Taxa de acerto e simulação de investimento do modelo Base da RNA com variáveis dependentes, utilizando dados estáticos.	68
Figura 15 – Taxa de acerto e simulação de investimento dos modelos Base e Base com atraso da RNA com variáveis dependentes, utilizando dados estáticos.	69
Figura 16 – Taxa de acerto e simulação de investimento dos modelos Base e Base com atraso da RNA com variáveis dependentes, utilizando dados dinâmicos calculados pela diferença relativa de 2 dias.	71
Figura 17 – Taxa de acerto e simulação de investimento dos modelos Base com atraso da RNA com variáveis dependentes, utilizando dados dinâmicos calculados pela diferença dos últimos 2 e 3 dias.	72
Figura 18 – Gráfico de dispersão da relação entre taxa de acerto e taxa de retorno do investimento simulado, obtido através dos experimentos realizados na RNA com Variáveis Exógenas e Indicadores.	73
Figura 19 – Taxa de acerto e simulação de investimento dos modelos Base e Base com Atraso da RNA com variáveis exógenas e indicadores, utilizando dados estáticos.	75

Figura 20 – Taxa de acerto e simulação de investimento dos modelos Base e Exógenos da RNA com variáveis dependentes e indicadores, utilizando dados dinâmicos calculados pela diferença relativa de 3 dias.	76
Figura 21 – Taxa de acerto e simulação de investimento dos modelos Base com da RNA com variáveis exógenas e indicadores, utilizando dados dinâmicos calculados pela diferença de 2 e de 3 dias.	77
Figura 22 – Taxa de acerto e simulação de investimento da fusão dos modelos Base da RNA com variáveis dependentes (1) e da RNA com variáveis exógenas e indicadores (2), utilizando dados dinâmicos calculados pela diferença de 2 dias.	79
Figura 23 – Taxa de acerto e simulação de investimento da fusão dos modelos Base com Atraso da RNA com variáveis dependentes (1) e da RNA com variáveis exógenas e indicadores (2), utilizando dados dinâmicos da diferença de 3 dias.	80
Figura 24 – Variação real das ações no período analisado, simulação de investimento com tendências aleatórias e com tendências geradas pelo modelo predictor final.	80

LISTA DE TABELAS

Tabela 1 – <i>Ranking</i> das 10 Ações com Maior Percentual de Participação em Volume Financeiro da Carteira de Ações do IBovespa.	55
Tabela 2 – Descrição das variáveis dependentes do ativo, em sua forma estática.	67
Tabela 3 – Descrição das variáveis dependentes do ativo, em sua forma dinâmica, geradas através do cálculo da diferença relativa.	70
Tabela 4 – Descrição das variáveis exógenas e de indicadores, em sua forma estática.	73
Tabela 5 – Descrição das variáveis exógenas e indicadores, em sua forma dinâmica, geradas através do cálculo da diferença relativa.	75

LISTA DE ABREVIATURAS E SIGLAS

CVM	Comissão de Valores Mobiliários
EMH	<i>Efficient-Market Hypothesis</i>
IBovespa	Índice Bovespa
ICON	Índice de Consumo
IEE	Índice de Energia Elétrica
IFNC	Índice Financeiro
IMAT	Índice Materiais Básicos
IMOB	Índice Imobiliário
INDX	Índice Industrial
IPO	<i>Initial Public Offering</i>
MACD	<i>Moving Average Convergence and Divergence</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MLP	<i>Multilayer Perceptron</i>
MME	Média Móvel Exponencial
MMS	Média Móvel Simples
MSE	<i>Mean Square Error</i>
NMSE	<i>Normalized Mean Square Error</i>
RMSE	<i>Root Mean Square Error</i>
RNA	Rede Neural Artificial
ROC	<i>Rate of Change</i>
SLG	<i>Sum of Losses and Gains</i>
UTIL	Índice Utilidade Pública

SUMÁRIO

1	INTRODUÇÃO	21
1.1	OBJETIVO	22
1.2	ESTRUTURA DO TRABALHO	22
2	MERCADO DE AÇÕES	23
2.1	INTRODUÇÃO	23
2.2	AÇÕES	24
2.3	COMERCIALIZAÇÃO DE AÇÕES	25
2.4	ANÁLISE DE AÇÕES	26
2.4.1	Análise Fundamentalista	27
2.4.2	Análise Técnica	29
2.5	CONSIDERAÇÕES FINAIS	33
3	SISTEMAS DE PREVISÃO DE SÉRIES TEMPORAIS FINANCEI- RAS	35
3.1	SÉRIES TEMPORAIS FINANCEIRAS	35
3.2	ARQUITETURA DOS PREVISORES	36
3.3	EXTRAÇÃO DAS CARACTERÍSTICAS	37
3.3.1	Seleção de Variáveis	37
3.3.2	Transformação das Variáveis	38
3.4	PREVISÃO	39
3.4.1	Redes Neurais Artificiais	40
3.4.1.1	Treinamento de Redes Neurais Artificiais	42
3.4.1.2	Função de Ativação	44
3.4.2	Máquinas de Vetor de Suporte	46
3.4.3	Outros Modelos Preditores	47
3.5	AVALIAÇÃO DOS RESULTADOS	48
3.5.1	Avaliação do Desempenho na Previsão de Valor Nominal	49
3.5.2	Avaliação do Desempenho na Previsão de Tendência	50
3.6	CONSIDERAÇÕES FINAIS	52
4	PROPOSTA DE SOLUÇÃO	53
4.1	ARQUITETURA DO SISTEMA	53
4.2	SELEÇÃO DAS SÉRIES FINANCEIRAS	54
4.3	SELEÇÃO DAS VARIÁVEIS DE ENTRADA	56
4.4	AVALIAÇÃO DO DESEMPENHO	57

5	DESENVOLVIMENTO E RESULTADOS	59
5.1	COLETA E PRÉ-PROCESSAMENTO DOS DADOS	59
5.2	CONFIGURAÇÃO DAS REDES NEURAIS	61
5.2.1	Arquitetura	61
5.2.2	Inicialização dos Pesos e Parâmetros das RNAs	62
5.2.3	Critério de Parada do Treinamento	62
5.2.4	Função de Ativação	64
5.2.5	Normalização dos Dados de Entrada	64
5.3	EXPERIMENTOS	65
5.3.1	RNA com Variáveis Dependentes	66
5.3.1.1	Variáveis de Entrada Estáticas	67
5.3.1.2	Variáveis de Entrada Dinâmicas	69
5.3.2	RNA com Variáveis Exógenas e Indicadores	72
5.3.2.1	Variáveis de Entrada Estáticas	73
5.3.2.2	Variáveis de Entrada Dinâmicas	75
5.4	FUSÃO	77
5.4.1	Fusão das Características	77
5.4.2	Fusão dos Sistemas Especialistas	78
6	CONSIDERAÇÕES FINAIS	81
	Referências	85

1 INTRODUÇÃO

As empresas podem abrir o seu capital aos investidores para captar recursos através da formação de ações. Ações são títulos negociáveis, representativos de uma fração do capital social de uma sociedade por ações (CASAGRANDE NETO; SOUSA; ROSSI, 2010). O dinheiro arrecadado com a venda das ações é revertido em benefício da empresa, sendo utilizado para financiar novos investimentos ou para quitar débitos por ela adquiridos. Em contrapartida, a empresa deve compensar os seus acionistas com a distribuição de dividendos e/ou bonificações, que podem ser em dinheiro ou através de novas ações. Qualquer remuneração de acionista deve ser proporcional ao número de ações que ele possui (KERR, 2011).

Segundo BM&FBovespa (2011), a Bolsa de Valores é o local que centraliza as negociações com ações. Além de organizar a compra e venda de produtos financeiros, a bolsa presta serviços de registro, compensação e liquidação, atuando como garantidora da liquidação financeira das operações realizadas em seus ambientes.

O preço pelo qual uma ação é negociada em bolsa, conforme Fortuna (2011), é fruto das condições de mercado (oferta e demanda) que refletem as condições estruturais e comportamentais da economia do País e específicas da empresa e de seu setor econômico. Portanto, como o preço de uma ação é composto de vários fatores, é necessário um grande conhecimento para realizar uma projeção, sendo esta uma tarefa complexa.

Para fazer uma projeção, é essencial ao investidor analisar a empresa em que está cogitando aplicar seus recursos. As formas de análise mais utilizadas para estimar o desempenho futuro das ações são a análise fundamentalista e a análise técnica (gráfica). Conforme Kerr (2011), a análise fundamentalista deve considerar elementos internos, como os balanços periódicos divulgados, e externos à empresa, como dados sobre o setor econômico de atuação e sobre o mercado acionário. Já a análise técnica parte da lógica que os preços passados influenciam nos preços futuros, não se preocupando com as causas das variações. Decisões são tomadas após analisar a representação gráfica destas variações.

Através da análise técnica, o investidor ou o analista de investimentos, pode utilizar técnicas ou indicadores, baseados em uma regressão nos dados históricos, para prever do comportamento futuro da série temporal. Stock e Watson (2004) definem uma série temporal como sendo um conjunto de dados coletados cronologicamente. As séries temporais financeiras se diferem muito entre si. Algumas têm comportamento estacionário, sendo possível prever o futuro olhando apenas para valores passados. Outras têm comportamento não-estacionário, neste caso os modelos estatísticos/computacionais buscam uma tendência na distribuição dos valores para prever os próximos valores da série.

No ramo de inteligência computacional existem técnicas e modelos capazes de realizar a previsão de séries temporais financeiras, como por exemplo as Redes Neurais Artificiais (YILDIZ; YALAMA; COSKUN, 2008) e Máquinas de Vetor de Suporte (TAY; CAO, 2001). Aplicações desenvolvidas com base nestas técnicas podem ser chamadas de sistemas especialistas, pois adquirem e armazenam conhecimentos, através de um treinamento, para depois desfrutar destes na solução de problemas similares. Sistemas especialistas usam conhecimentos específicos do domínio de um problema para conseguir obter um desempenho de qualidade especialista em determinada área de atuação (LUGER, 2013).

A previsão de algumas séries temporais financeiras pode ser muito complexa, devido ao seu comportamento não-estacionário. A resolução de uma tarefa computacional complexa, segundo Haykin (2001), pode ser obtida dividindo-a em um número de tarefas computacionais simples. Estas sub-tarefas são direcionadas aos sistemas especialistas. Os resultados obtidos por cada especialista são combinados para chegar a uma decisão global, que supostamente é melhor que aquela alcançável por qualquer especialista atuando isoladamente.

1.1 OBJETIVO

Neste contexto, o presente trabalho tem como objetivo geral desenvolver uma aplicação, baseada na combinação de sistemas especialistas, que faça a previsão do comportamento futuro do preço de ações de algumas empresas com grande influência na economia nacional.

1.2 ESTRUTURA DO TRABALHO

Este trabalho está estruturado de forma que neste capítulo é apresentada uma introdução ao assunto previsão de séries temporais financeiras. No [Capítulo 2](#) são abordados assuntos referentes ao mercado de ações como a comercialização e formas de análise de ações. No [Capítulo 3](#) são tratadas questões relevantes à modelagem de sistemas de previsão de séries temporais financeiras. No [Capítulo 4](#) é apresentada a proposta de solução desenvolvida. Por sua vez, o [Capítulo 5](#) apresenta as etapas de desenvolvimento do sistema, bem como os resultados obtidos na previsão das séries temporais financeiras selecionadas. Por fim, no [Capítulo 6](#) são apresentadas algumas considerações sobre o trabalho.

2 MERCADO DE AÇÕES

Este capítulo introduz conceitos sobre o mercado de ações. Na [seção 2.1](#) é explicada, de forma sucinta, a estrutura do mercado financeiro brasileiro. Na [seção 2.2](#) são apresentadas definições importantes sobre as ações e o mercado de ações. A [seção 2.3](#) aborda a comercialização das ações, tratando de assuntos como bolsas de valores e mercados primário e secundário. As formas de análise de ações utilizadas tradicionalmente por analistas são assunto para a [seção 2.4](#). Na [seção 2.5](#) são apresentadas algumas considerações sobre o capítulo.

2.1 INTRODUÇÃO

O mercado financeiro é um ambiente no qual ocorre o relacionamento entre os atores que possuem recursos monetários (agentes superavitários) e os que precisam de recursos para realizar investimentos ou pagar dívidas (agentes deficitários) ([PEREIRA, 2013](#)). O mercado financeiro está dividido em 4 segmentos, segundo [Pereira \(2013\)](#): mercado monetário, mercado de crédito, mercado de câmbio e mercado de capitais.

O mercado monetário é o espaço no qual são realizadas operações de curto e curtíssimo prazo ([KERR, 2011](#)). Tem como objetivo controlar a liquidez dos meios de pagamento da economia, através de operações de mercado aberto, de redesconto e depósitos compulsórios.

O mercado de crédito é o espaço de atuação de instituições financeiras e não financeiras, que fazem a intermediação de recursos, de curto e médio prazo, entre os agentes superavitários e deficitários ([PEREIRA, 2013](#)). Este mercado também é denominado mercado bancário. Destina-se basicamente a suprir as necessidades de capitalização de empresas e financiamento de consumo de indivíduos.

O mercado de câmbio compreende as operações de compra e venda de moedas estrangeiras, realizadas em instituições que possuem autorização do Banco Central ([FORTUNA, 2011](#)). As operações relativas ao recebimento, pagamento ou transferência para e do exterior via cartão de crédito ou débito internacional, ou transferências entre contas internacionais, também incluem-se no mercado cambial.

O mercado de capitais é o sistema de distribuição de valores mobiliários cujo objetivo é proporcionar liquidez aos títulos emitidos pelas empresas e viabilizar o seu processo de capitalização ([PEREIRA, 2013](#)). Compreende as operações de prazo médio, longo ou indeterminado (debêntures) e operações de prazo indeterminado (ações).

2.2 AÇÕES

Uma ação representa a menor parcela do capital social de uma empresa organizada sob a forma de sociedade anônima (KERR, 2011). As sociedades anônimas podem ser de capital aberto (ações livremente negociadas no mercado) ou fechado (venda e emissão de ações de modo particular). Ao adquirir ações de uma empresa, o investidor se torna um de seus proprietários, com direito a participar dos resultados alcançados pela empresa proporcionalmente à quantidade de ações que ele possui.

O mercado de ações¹ e valores mobiliários em geral não oferecem ao investidor rentabilidade garantida (BM&FBOVESPA, 2011). Por não oferecer garantia de retorno, devem ser considerados investimentos de risco, pois o investidor, que tem como objetivo ganhar dinheiro, pode ter seu capital reduzido ao investir em produtos deste mercado.

As ações podem ser de dois tipos: ordinárias ou preferenciais (FORTUNA, 2011). Ações ordinárias conferem aos seus detentores o direito a voto nas assembleias. As ações preferenciais conferem aos seus detentores o direito de preferência sobre os dividendos a serem distribuídos entre os acionistas e de retorno do capital investido em caso de dissolução da empresa.

Entende-se por dividendos, a distribuição de parte dos lucros da empresa aos seus acionistas (FORTUNA, 2011). A empresa deve distribuir pelo menos 25% do lucro líquido do exercício entre os acionistas. Além dos dividendos, os acionistas podem receber bonificações que ocorrem em função de aumento do capital das empresas por incorporação de reservas (KERR, 2011). A bonificação pode ser concedida pela distribuição de novas ações ou pela distribuição de dinheiro, sempre em quantidade proporcional à participação do acionista na empresa.

Outro benefício importante do acionista, os direitos de subscrição, ocorrem quando a empresa decide aumentar o seu capital através da emissão de novas ações (KERR, 2011). A empresa tem a obrigação de oferecer as novas ações aos atuais acionistas, por um preço inferior ao da cotação atual. Caso o acionista decida não exercer o direito de subscrição ele terá o percentual de participação na empresa diluído pela emissão destas novas ações.

A subscrição de novas ações aumenta o número de ações da empresa mas não altera o seu valor unitário, o que aumenta é o valor do capital social da empresa. Porém, existem outros dois eventos que alteram o número de ações sem alterar o capital da empresa, modificando o preço de cada ação no mercado: o desdobramento de ações (*split*) e o agrupamento de ações (*inplit*) (FORTUNA, 2011).

O desdobramento de ações consiste na distribuição gratuita de novas ações aos acionistas, pela diluição do capital em número maior de ações. Este ato tem como objetivo

¹ Neste trabalho, entende-se por mercado de ações o mercado à vista de ações, onde a compra e venda de ações ocorre com os preços cotados no momento da emissão da ordem de compra ou venda.

principal dar liquidez aos títulos da empresa no mercado (FORTUNA, 2011). Geralmente ocorre quando a ação da empresa está cotada a preços muito elevados (KERR, 2011). Se, por exemplo, for emitida uma nova ação para cada ação existente no mercado, diz-se que houve um desdobramento de 1:2 (lê-se um para dois), onde cada ação passará a valer 50% do valor antigo, e cada acionista terá o dobro das ações que detinha.

O agrupamento de ações é o inverso do desdobramento. Geralmente ocorre quando a ação está cotada a preços muito baixos, o que pode aumentar a volatilidade e causar impressões ruins em potenciais acionistas (KERR, 2011). Assim como no desdobramento, o agrupamento não altera o valor que detém cada acionista nem o capital da empresa, apenas modifica o número de ações que ele detém e o preço unitário de cada ação.

2.3 COMERCIALIZAÇÃO DE AÇÕES

A abertura de capital de uma empresa é o processo de transformação de uma empresa em companhia aberta (CASAGRANDE NETO; SOUSA; ROSSI, 2010). A Oferta Pública Inicial (do Inglês *Initial Public Offering* - IPO) ocorre no momento da abertura de capital social da empresa, quando as ações da empresa são colocadas a venda pela primeira vez. A IPO requer um prévio registro do emissor (empresa) na Comissão de Valores Mobiliários (CVM), que é uma autarquia vinculada ao Ministério da Fazenda responsável por disciplinar e fiscalizar o mercado de valores mobiliários. Os recursos obtidos através da IPO são revertidos em benefício da empresa e utilizados para financiar novos investimentos ou para quitar dívidas.

O mercado de ações pode ser classificado em mercado primário e mercado secundário (FORTUNA, 2011). No mercado primário são comercializadas as ações emitidas diretamente pela empresa, através de ofertas públicas. Já no mercado secundário as ações são negociadas entre os investidores nas bolsas de valores, a empresa não recebe valores monetários com estas negociações. O mercado secundário é o responsável por dar liquidez ao mercado primário, pois dificilmente um investidor se tornaria sócio de uma empresa se não conseguisse sair da sociedade facilmente.

As bolsas de valores podem ser definidas como mercados de leilões organizados, onde são comercializados produtos financeiros, dentre eles as ações (KERR, 2011). No Brasil a única bolsa de valores é a BM&FBovespa, que surgiu da fusão, em 2007, de duas grandes bolsas de São Paulo: a Bolsa de Valores de São Paulo (Bovespa) e a Bolsa de Mercadorias e Futuros (BM&F). Até o final dos anos 90 as negociações ocorriam fisicamente nos prédios das bolsas de valores, através do pregão viva-voz.

Com o advento da Internet, no final dos anos 90, foi instituído o pregão eletrônico, onde os investidores passam a operar eletronicamente através dos *sites* das corretoras de valores (FORTUNA, 2011). As corretoras de valores têm duas principais funções: executar

as ordens de compra ou venda do investidor e auxiliá-lo na construção de sua carteira de ações e na análise prévia das empresas (BM&FBOVESPA, 2011).

Inicialmente, o indivíduo que deseja investir em ações deve se cadastrar em uma corretora de valores. Esse cadastro criará uma conta do investidor junto à corretora (como se fosse uma conta bancária), o investidor deve então transferir recursos monetários para essa conta (BM&FBOVESPA, 2011). Após este processo, o investidor pode começar a operar através do *Home Broker*, um sistema que permite enviar ordens de compra e venda diretamente de seu computador via Internet (KERR, 2011). Estas ordens de compra e venda são enviadas à corretora, que oficializará as ordens junto a Bolsa de Valores. A corretora cobra uma taxa de corretagem para cada ordem de compra ou venda emitida pelo investidor.

As bolsas de valores disponibilizam índices de desempenho médio do mercado, utilizados como referência direta para a análise do comportamento individual das ações de um grupo específico (FORTUNA, 2011). O Índice Bovespa (IBovespa) é o mais importante indicador de desempenho médio do mercado de ações brasileiro. Sua carteira teórica é integrada pelas ações que, em conjunto, representam aproximadamente 85% dos negócios e do volume comercializado. A carteira de ações do IBovespa é recalculada a cada 4 meses, com base nos dados de comercialização das ações do período de vigência das últimas 3 carteiras (1 ano) (BM&FBOVESPA, 2014).

Os índices setoriais, também fornecidos pela Bolsa, tem por objetivo oferecer uma visão segmentada do mercado acionário, medindo o comportamento médio das ações de empresas representativas nos setores em que elas atuam (BM&FBOVESPA, 2014). Os principais índices setoriais da BM&FBovespa são: IEE (Energia Elétrica), INDX (Industrial), ICON (Consumo), IMOB (Imobiliário), IFNC (Financeiro), IMAT (Materiais Básicos) e UTIL (Utilidade Pública).

2.4 ANÁLISE DE AÇÕES

Segundo Fortuna (2011), o preço de uma ação em bolsa é fruto das condições de mercado (oferta e demanda) que refletem as condições estruturais e de comportamento da economia do País e específicas da empresa e de seu setor econômico. O preço das ações traduz as expectativas dos agentes econômicos em relação às perspectivas do País e, por consequência, aos destinos das empresas de capital aberto que atuam naquele mercado.

Os analistas do mercado de ações costumam dizer que qualquer ação possui um valor de mercado e um valor intrínseco (KERR, 2011). O mercado nem sempre estima corretamente o preço da ação, precificando um valor diferente do real (intrínseco), que pode ser maior ou menor que o valor de mercado. Quando o mercado precifica de forma correta uma determinada ação, o valor intrínseco é igual ao valor de mercado. Esta lógica

pressupõe que o mercado não é eficiente ou não está eficiente no momento da determinação de um preço diferente de seu valor de mercado.

Segundo Kerr (2011), a hipótese do mercado eficiente (do Inglês *Efficient-Market Hypothesis* - EMH) diz que um mercado é eficiente quando ajusta seus preços rapidamente à chegada de novas informações. O preço de uma ação deve refletir todas as informações disponíveis no momento, em um âmbito público e privado.

Para os investidores, conhecer previamente o preço futuro das ações pode significar a diferença entre o sucesso e o fracasso nos investimentos em produtos financeiros (RIBEIRO; GOLDSCHMIDT; CHOREN, 2009). Para evitar possíveis perdas, o investidor procura embasar suas decisões através do uso de métodos de análise de ações, objetivando maximizar seus ganhos. Esta análise tradicionalmente pode ser feita de duas formas: análise fundamentalista e análise técnica.

2.4.1 Análise Fundamentalista

Para Fortuna (2011), a escola fundamentalista baseia-se nos resultados setoriais e específicos de cada empresa, dentro do contexto da economia nacional e internacional. Essa análise é fundamental na escolha da empresa cuja ação deve ser adquirida, bem como na determinação de qual ação deve ser vendida em um dado intervalo de tempo.

Sua premissa básica é que o valor justo para uma ação de uma empresa está diretamente relacionado à sua capacidade de gerar lucros (SACHETIM, 2006). Os principais indicadores de análise fundamentalista que auxiliam os investidores na decisão da aquisição ou venda de ações são (FORTUNA, 2011):

- Preço da ação no mercado: se refere à cotação da ação em bolsa;
- Lucro por ação: corresponde ao lucro líquido anual da ação;
- Índice preço da ação no mercado/lucro por ação (P/L): este índice dá uma ideia do tempo (em anos) necessário para recuperar o investimento realizado, imaginando que a empresa manterá o mesmo lucro nos anos seguintes. Se o P/L é baixo indica que a ação está custando pouco em relação ao lucro que ela representa;
- Índice preço da ação no mercado/valor patrimonial da empresa (P/VP): indica quanto o mercado está disposto a pagar por uma ação em relação ao valor patrimonial da empresa. $P/VP < 1$ indica que o preço da ação está abaixo do valor patrimonial por ação da empresa;
- Índice *dividend yield*: corresponde ao dividendo pago pela ação/preço da ação no mercado. Quanto maior o *dividend yield* maiores são as chances do investidor recuperar seus investimentos através do recebimento de dividendos;

- Índice *pay out*: equivale ao dividendo pago pela ação/lucro por ação;
- Índice preço da ação no mercado/fluxo de caixa operacional: indica quantos anos são necessários para a empresa gerar caixa equivalente ao valor total de suas ações;
- Retorno sobre o Patrimônio Líquido (ROE): corresponde ao lucro líquido dividido pelo patrimônio líquido.

A função de um analista fundamentalista é determinar os porquês das tendências de alta ou de baixa das ações (SACHETIM, 2006). Além de observar os índices fundamentalistas das ações, o analista deve realizar uma cuidadosa análise dos procedimentos operacionais da empresa, como balanços e demonstrativos financeiros. Busca identificar as perspectivas futuras daquela empresa no mercado.

A lógica por trás da análise fundamentalista é que, sendo as ações parte de uma empresa, é possível avaliar a empresa e descobrir como será seu desempenho futuro (KERR, 2011). A partir disso é possível estimar o preço futuro de suas ações. A análise fundamentalista está diretamente ligada à EMH por acreditar que através das informações disponíveis é possível estimar corretamente o preço das ações.

A comparação dos desempenhos de empresas que atuam em um mesmo setor econômico pode destacar aquelas que têm mais capacidade de geração de lucros. Neste contexto, os procedimentos que um analista fundamentalista deve observar são os seguintes (SACHETIM, 2006):

- Análise do setor em que a empresa atua: o analista deve buscar informações referentes ao setor de atuação para que possam ser criados cenários futuros das empresas daquele setor, com relação a disponibilidade de matéria-prima e políticas governamentais para o setor, por exemplo;
- Análise da estrutura organizacional da empresa: visa identificar quais são os planos de investimentos, de marketing, de melhorias tecnológicas e qual é a eficiência administrativa da corporação;
- Análise dos demonstrativos financeiros: esta análise permite fazer projeções de resultados futuros, onde devem ser analisados os dados da empresa e comparados com os dados de outras empresas do mesmo setor.

Em resumo, o analista fundamentalista deve analisar o setor de atuação e a situação econômica da empresa, além dos indicadores fundamentais das ações, para tomar uma decisão com potencial de bons retornos financeiros. Portanto, a análise fundamentalista não é uma tarefa fácil, pois além de exigir a análise de uma grande conjuntura de dados, o analista deve saber interpretar corretamente estas informações. Normalmente analistas fundamentalistas se especializam em uma determinada empresa ou em um setor econômico.

2.4.2 Análise Técnica

Para [Murphy \(1999\)](#), a análise técnica é o estudo da ação no mercado, por meio do uso de gráficos, com o propósito de prever tendências futuras dos preços. [Murphy \(1999\)](#) também afirma que o grafista (analista que observa apenas os gráficos para tomar suas decisões) sabe que existem razões para a ocorrência de alterações nos preços, mas acredita que não é necessário saber quais são estas razões para prever o movimento do mercado.

A análise técnica se baseia em três premissas ([MURPHY, 1999](#)): o mercado assimila fatos políticos, econômicos e psicológicos na formação do preço das ações; os preços se movem em tendências; preços futuros serão uma repetição dos preços passados. A primeira premissa indica que a análise fundamentalista das ações não é suficiente para prever os preços, uma vez que nos preços estão refletidos fatos políticos e psicológicos, que por sua natureza são difíceis de mensurar. A segunda e terceira premissa indicam que é possível fazer a previsão de valores futuros identificando a tendência (de alta ou de baixa) e analisando os dados passados. Analistas técnicos não acreditam na EMH.

Segundo [Fortuna \(2011\)](#), uma boa análise técnica dispensa a pesquisa sobre os fundamentos da empresa, pois os gráficos são a soma de todos os conhecimentos, esperanças e expectativas sobre uma determinada ação. À medida que os fundamentos de uma empresa passam a ser fatos concretos eles automaticamente são incorporados aos gráficos, pois a movimentação do preço das ações em mercado reflete os acontecimentos.

A análise técnica tem como base os volumes de negociação e os preços pelos quais foram negociadas as ações em pregões anteriores ([FORTUNA, 2011](#)). Sobre estes dados são aplicados indicadores técnicos, que objetivam identificar a situação atual do ativo e prever o seu futuro. Os indicadores podem ser divididos em dois grupos: os rastreadores de tendência e os osciladores ([SACHETIM, 2006](#)).

Os rastreadores de tendência foram criados primeiramente para identificar quando um mercado está subindo ou descendo ([SACHETIM, 2006](#)). O propósito do método de identificação de tendências é olhar o passado para encontrar a direção atual dos preços. Esses indicadores são confiáveis durante períodos de tendência, que ocorrem em somente 30% do mercado, e não identificam o início ou fim de um período de tendência ([MURPHY, 1999](#)). Os principais indicadores técnicos utilizados na identificação de tendências utilizam médias móveis na sua formação. A média móvel mostra o valor médio dos dados em determinado período e é utilizada para suavizar os ruídos das flutuações dos preços de curto prazo, de maneira que seja mais fácil identificar as tendências ([VIDOTTO; MIGLIATO; ZAMBON, 2009](#)). Três dos principais índices rastreadores de tendências são ([SACHETIM, 2006](#)):

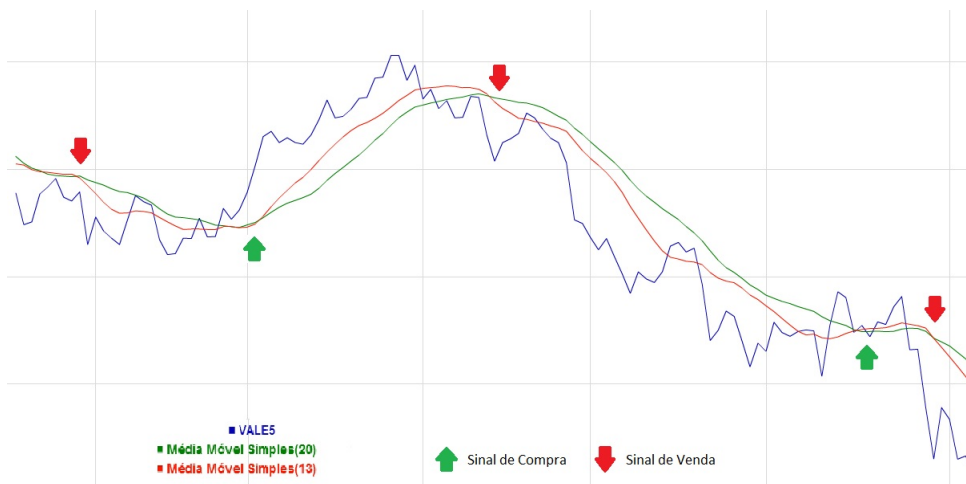
- Intersecção de Médias Móveis: consiste na criação de duas médias móveis simples (MMS), uma de longo e outra de curto prazo, que são exibidas em um mesmo gráfico

para a análise. A média móvel simples pode ser obtida através de

$$MMS = \frac{1}{n} \sum_{d=t-n}^n v_{t-d} \quad (2.1)$$

onde t representa o dia atual, n representa o período de cálculo da MMS e d é o número de dias de regressão na série temporal em relação ao dia atual. A MMS de maior período identifica a tendência dos preços a longo prazo, a MMS de menor período identifica a tendência a curto prazo. Quando a linha da média de curto prazo está acima da de longo prazo identifica-se uma tendência de alta dos preços. Este indicador também pode gerar sinais de compra e venda. Os sinais de compra ou venda das ações ocorrem quando as linhas da média móvel de longo e curto prazo se cruzam (SAFFI, 2003). Compra, quando a média de curto prazo cruza a outra, se posicionando abaixo da linha de longo prazo; e venda, quando a média de curto prazo corta a de longo se posicionando acima da linha de longo prazo. Este comportamento é ilustrado na Figura 1, que exemplifica a combinação mais utilizada, 13 dias para média de curto prazo (linha vermelha) e 20 dias para a média de longo prazo (linha verde). Os sinais de compra são representados pela seta verde e os sinais de venda pela seta vermelha.

Figura 1: Intersecção de Médias Móveis - ação VALE5 da empresa Vale.



Fonte: (Bússola do Investidor, 2014) (Adaptado)

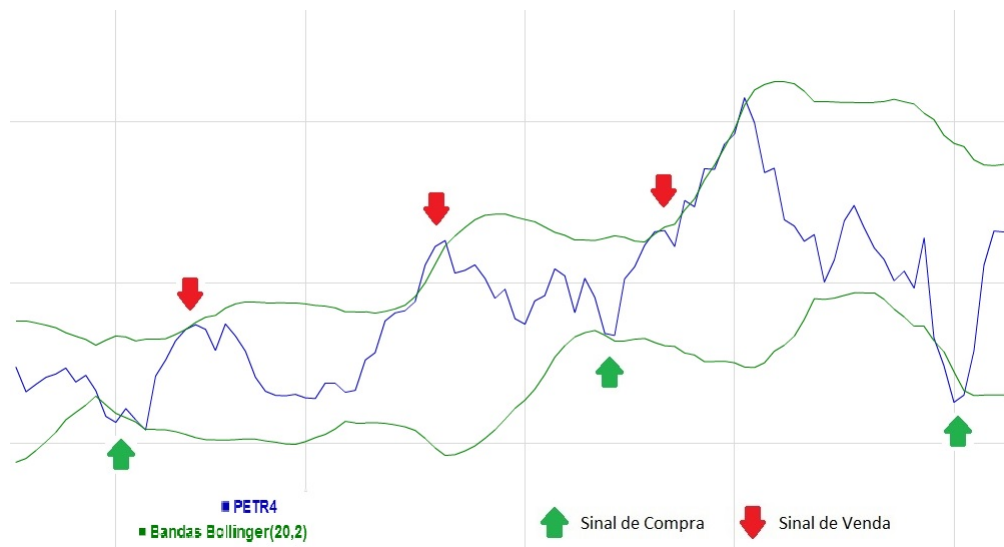
- Bandas de Bollinger: do Inglês *Bollinger Bands*, também utilizam uma média móvel simples, que vem acompanhada de duas bandas, uma por cima e outra por baixo da média. A banda central é formada através do cálculo de média móvel simples dos últimos 20 dias. As bandas superior e inferior são formadas pelo acréscimo e decréscimo, respectivamente, de duas unidades do desvio-padrão dos últimos 20 dias.

O desvio-padrão pode ser obtido através de

$$DP = \frac{1}{n-1} \sum_{i=1}^n (v_i - MMS)^2 \quad (2.2)$$

onde n é o período para o qual foi calculada a MMS, v_i é o valor da amostra no dia i que pertence período da média. Este indicador também tem como objetivo a determinação de tendências de alta ou de baixa. Quando a linha do preço de fechamento está próximo da banda inferior indica uma tendência de baixa do valor da ação. Quando está próximo da banda superior a tendência é de alta. O cruzamento da linha do preço de fechamento com a banda inferior gera um sinal de compra; o cruzamento com a banda superior gera um sinal de venda. A Figura 2 exemplifica o uso das Bandas de Bollinger (neste gráfico a banda central não é exibida) na geração de sinais de compra (seta verde) e venda (seta vermelha). Percebe-se que o indicador não tem uma precisão exata sobre o momento de realizar a compra ou venda dos títulos. Mas, na maioria das vezes a sugestão de compra ou venda ocorre em momento adequado.

Figura 2: Bandas de Bollinger - ação PETR4 da empresa Petrobrás.



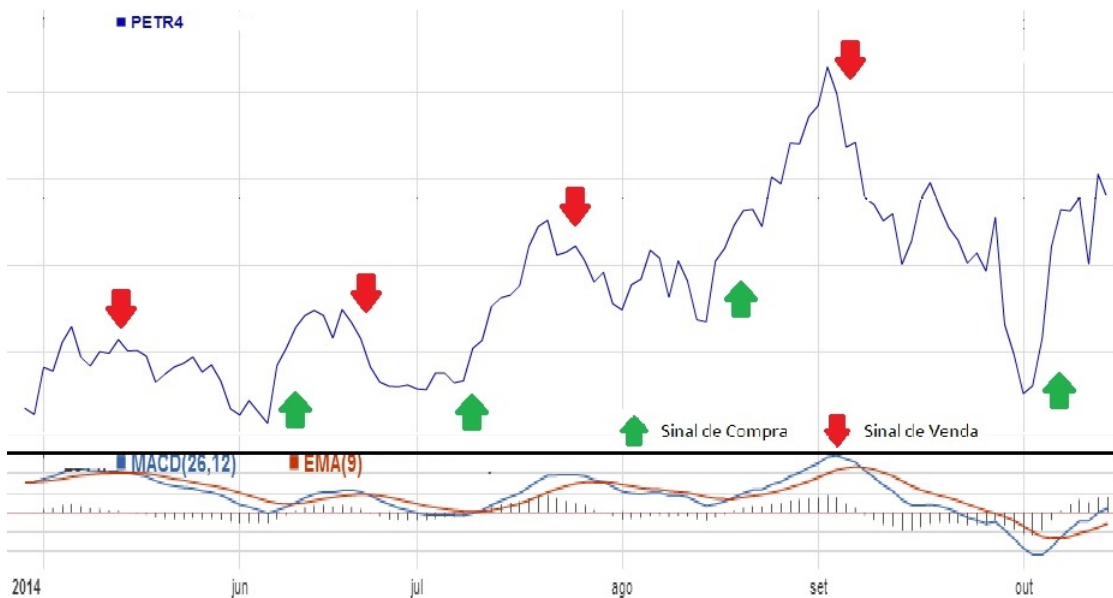
Fonte: (Bússola do Investidor, 2014) (Adaptado)

- Convergência e Divergência da Média Móvel: O indicador Convergência e Divergência da Média Móvel, do Inglês *Moving Average Convergence and Divergence* - MACD, caracteriza-se pela combinação de três médias móveis exponenciais (MME). A MME pode ser obtida através de

$$MME_d = (p_d - MME_{d-1}) * k + MME_{d-1} \quad (2.3)$$

onde p_d é o preço do dia atual d , MME_{d-1} ² é a MME calculada para o dia anterior e k calculado por $k = \frac{2}{n+1}$, onde n é o número de dias escolhido para o cálculo da MME. O MACD é composto de duas linhas: a linha de sinal (que corresponde a MME de 9 dias) e linha MACD (que corresponde a diferença entre a MME de 26 dias e a MME de 12 dias). O cruzamento da linha MACD de baixo para cima da linha de sinal equivale a um sinal de compra; já o cruzamento da linha MACD de cima para baixo da linha de sinal equivale a um sinal de venda. A Figura 3 exemplifica o uso do indicador MACD na geração de sinais de compra (seta verde) e venda (seta vermelha).

Figura 3: Convergência e Divergência da Média Móvel (MACD) - ação PETR4 da empresa Petrobrás.



Fonte: (Bússola do Investidor, 2014) (Adaptado)

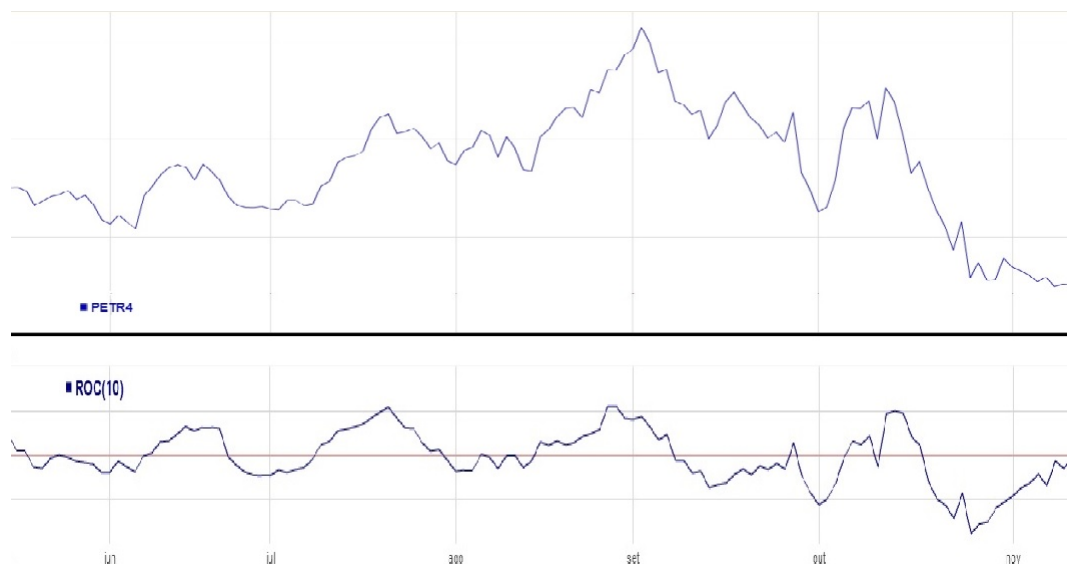
Os indicadores de oscilação, ou osciladores, são denominados desta forma pois são eficientes em momentos em que o mercado não está com uma tendência definida, onde a maioria dos sistemas de acompanhamento de tendências não funcionam bem (MURPHY, 1999). Os osciladores são considerados indicadores secundários e devem ser utilizados em paralelo com indicadores de rastreamento de tendências. Dentre os osciladores, o indicador taxa de variação, do Inglês *Rate of Change* - ROC, é um dos mais utilizados. O ROC é obtido através de uma relação entre o preço de fechamento mais recente e o preço de fechamento de x dias atrás. Pode ser calculado como

$$ROC_x = \frac{p}{p_x} \quad (2.4)$$

² A primeira MME é calculada através de uma média simples dos últimos n dias, onde n é o período para o qual deseja-se calcular a MME.

onde p é o preço de fechamento mais recente, p_x é o preço de fechamento de x dias atrás. A Figura 4 representa a utilização do indicador ROC, neste caso representando a taxa de variação do preço nos últimos 10 dias.

Figura 4: ROC_{10} - ação PETR4 da empresa Petrobrás.



Fonte: (Bússola do Investidor, 2014) (Adaptado)

Além dos indicadores citados, existem vários outros que podem ser utilizados na análise técnica de ações. São exemplos de outros indicadores: indicador de força relativa (IFR), estocástico lento, oscilador de Chaikin (CHO), médias móveis exponenciais, médias móveis modificadas (MMA), SAR Parabólico, índice direcional médio (DMI), entre outros (Bússola do Investidor, 2014).

2.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram introduzidos conceitos de fundamental importância para o decorrer deste trabalho. Destacam-se as formas de análise de ações e as diferenças entre as duas escolas: fundamentalista e técnica.

A análise fundamentalista apresenta-se como de difícil compreensão aos investidores. Para realizar este tipo de análise o investidor deve possuir um grande conhecimento em contabilidade (para realizar a análise dos dados contábeis da empresa) e ter uma boa visão macroeconômica (para projetar o futuro do setor de atuação da empresa e do mercado em que a ação está inserida). Com a utilização da análise fundamentalista o analista tem uma visão de médio a longo prazo do futuro da empresa em questão.

A análise técnica, parte do princípio que os preços passados influenciam nos preços futuros. É sobre este pilar da análise técnica que se fundamenta a tentativa de realizar

a previsão futura das ações, que é o objetivo deste trabalho. A análise técnica é mais empregada e apresenta melhores resultados quando o objetivo é realizar uma análise de ações a curto prazo.

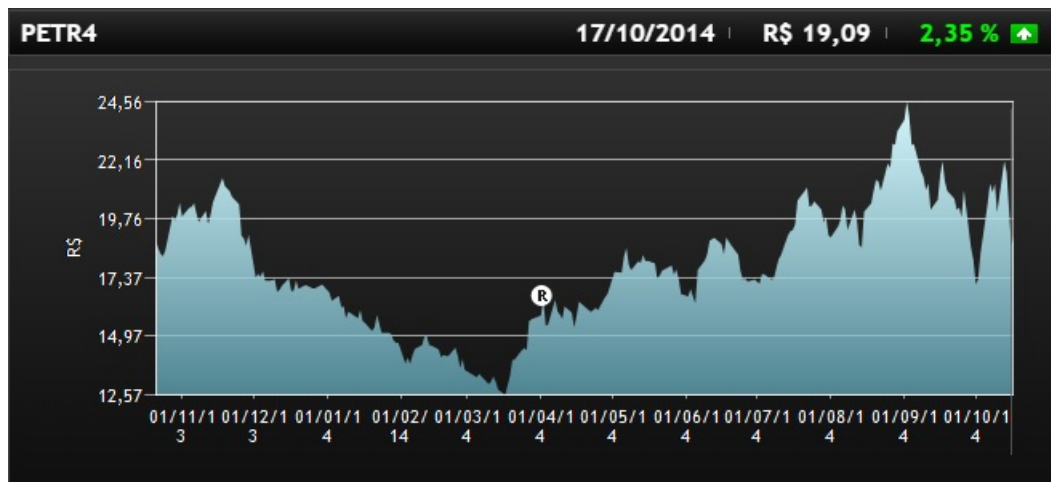
3 SISTEMAS DE PREVISÃO DE SÉRIES TEMPORAIS FINANCEIRAS

Neste capítulo são apresentadas as principais características de sistemas previsores de séries temporais financeiras. A [seção 3.1](#) introduz o conceito de série temporal. Na [seção 3.2](#) é abordada a arquitetura de sistemas reconhedores de padrões aplicados na previsão de séries financeiras. A seleção de variáveis para uso na previsão e extração de suas características é tratada na [seção 3.3](#). As metodologias utilizadas na previsão e as formas de avaliação dos resultados alcançados são abordadas na [seção 3.4](#) e [seção 3.5](#), respectivamente. Na [seção 3.6](#) são apresentadas algumas considerações sobre o capítulo.

3.1 SÉRIES TEMPORAIS FINANCEIRAS

Série temporal é um conjunto de observações ordenadas no tempo ([MORETTIN; TOLOI, 1987](#)). Uma série temporal pode ser representada por um gráfico em que o tempo é mostrado no eixo horizontal (eixo x), enquanto a variável em estudo tem seus valores mostrados no eixo vertical (eixo y). A [Figura 5](#) mostra a representação gráfica de uma série temporal. Outros exemplos de séries temporais da área financeira são os valores: de cotação do dólar, de índices setoriais ou de índices de mercado como o IBovespa.

Figura 5: Série Temporal da ação PETR4 da Empresa Petrobrás, de 22 de outubro de 2013 a 17 de outubro de 2014.



Fonte: ([BM&FBOVESPA, 2014](#))

A série temporal pode ser representada como ([CASTANHEIRA, 2008](#))

$$y = f(t) \quad (3.1)$$

onde y é a variável dependente, $f(.)$ é a função que relaciona o valor de y com a data de referência e t é a data de referência. A previsão de séries temporais consiste em encontrar a função $f(.)$.

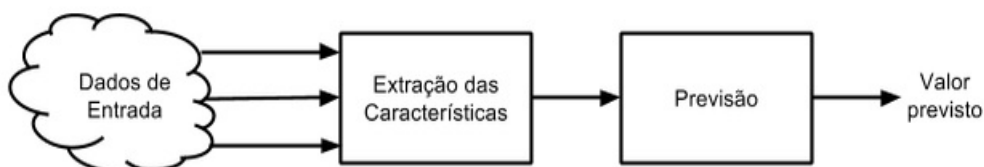
A análise técnica de ações, como visto no [Capítulo 2](#), parte do princípio que o movimento dos preços passados influencia em preços futuros, havendo um comportamento repetitivo ao longo do tempo. Dessa forma, apesar das condições econômicas e políticas serem diferentes, a variação dos preços de ativos financeiros se comporta de maneira similar a variações que ocorreram em algum momento no passado.

3.2 ARQUITETURA DOS PREVISORES

Os sistemas previsores de séries financeiras podem ser tomados como um problema de reconhecimento de padrões. [Theodoridis e Koutroumbas \(2003\)](#) definem o reconhecimento de padrões como uma disciplina científica que tem por objetivo classificar objetos em um número de categorias ou classes. Estes objetos, dependendo da aplicação, podem ser imagens, sinais ou números que possam de alguma forma ser classificados. [Haykin \(2001\)](#) complementa definindo o reconhecimento de padrões como sendo um processo pelo qual um sinal recebido é atribuído a uma classe dentre um número predeterminado de categorias (classes). O objetivo do reconhecimento de padrões é esclarecer os mecanismos complexos do processo de tomada de decisão e automatizar essas funções utilizando computadores ([FUKUNAGA, 1990](#)).

O processo de criação de sistemas de previsão de preço/tendência de produtos financeiros pode ser dividido em duas etapas, a extração das características e a previsão em si, como ilustra a [Figura 6](#). A etapa de extração de características consiste na seleção das variáveis e na transformação sobre elas realizada visando realçar as suas características ou modificá-las de modo a facilitar o processo de classificação/previsão. Estas variáveis servirão de base para que o modelo predictor, construído com técnicas computacionais, possa "aprender" o padrão existente nos dados e fazer a previsão dos próximos valores.

Figura 6: Arquitetura genérica de sistemas de previsão.



Fonte: desenvolvido pelo autor

3.3 EXTRAÇÃO DAS CARACTERÍSTICAS

A classificação de objetos em determinada classe somente é possível se os objetos das classes se diferem uns dos outros por determinadas peculiaridades (DOUGHERTY, 2013). Em um processo de classificação, um objeto tem suas características extraídas e passadas para um classificador. Este deve avaliar os valores e tomar uma decisão indicando a classe que o objeto deve ser atribuído.

O processo de extração de características pode ser dividido em duas etapas: a seleção de variáveis e a transformação sobre elas realizada a fim de evidenciar suas características. Através da transformação de uma variável também é possível criar outro dado que seja representativo no processo de classificação.

3.3.1 Seleção de Variáveis

A seleção das variáveis é uma tarefa muito importante e afeta fortemente o projeto de um classificador (FUKUNAGA, 1990). Se as variáveis apresentam diferenças significativas em suas características de uma classe para outra, o classificador pode ser concebido facilmente e apresentar ótimo desempenho.

As variáveis selecionadas para o treinamento do modelo predictor podem ser separadas em dois grupos: variáveis dependentes do ativo que se busca fazer a previsão e variáveis exógenas. As variáveis exógenas são independentes daquela que se busca prever o futuro, não possuem uma ligação direta com o ativo a ser previsto.

O grupo de variáveis dependentes do ativo é composto das séries históricas de cotações relacionadas com o ativo a ser previsto. Estas séries correspondem a um resumo diário das negociações de um ativo em um dia determinado e são fornecidas pelas bolsas de valores. No total, as bolsas de valores fornecem 5 valores diários que resumem a negociação de um ativo em um determinado dia: preço de fechamento, preço de abertura, preço mais alto, preço mais baixo e volume negociado. A série do preço de fechamento, que é a série que deseja-se prever o futuro, sempre está entre os dados selecionados para o treinamento do modelo de previsão. Em alguns casos, os dados históricos do preço de fechamento são os únicos dados utilizados na modelagem do sistema predictor (TAY; CAO, 2001) (MACIEL; BALLINI, 2009) (CHARKHA, 2008) (ZHANG; JIANG; LI, 2004) (LI; XIONG, 2005) (FARIA et al., 2008). As outras 4 séries de cotações históricas dependentes (preço de abertura, preço mais alto, preço mais baixo e volume de negociação) são frequentemente selecionadas para se integrarem ao modelo de previsão (MACEDO et al., 2013) (NETO; CAVALCANTI et al., 2008) (ABRAHAM; PHILIP; SARATCHANDRAN, 2004) (YILDIZ; YALAMA; COSKUN, 2008) (WU; LU, 2012) (NAEINI; TAREMIAN; HASHEMI, 2010) (KHAN; BANDOPADHYAYA; SHARMA, 2008).

No segundo grupo estão as variáveis exógenas, que são escolhidas por apresentarem

alguma possível relação de influência na formação do preço do ativo. Uma das variáveis exógenas frequentemente selecionada é a taxa de câmbio de moedas, pois moedas estrangeiras podem influenciar no preço de determinados ativos financeiros. Neto, Cavalcanti et al. (2008) selecionaram a cotação do Dólar em Reais e a cotação do Euro em Reais, para auxiliar na previsão do preço da ação da Petrobrás. Da mesma forma, Yildiz, Yalama e Coskun (2008) empregaram em seu modelo predictor do índice da Bolsa de Valores de Istambul a cotação do Dólar em relação à moeda local. Ainda no que se refere a variáveis exógenas selecionadas, os índices de bolsas de valores por vezes também são utilizados no processo de previsão de ativos financeiros. Esses índices, conforme visto no Capítulo 2 são indicadores do desempenho médio de um conjunto de ações, de um país ou de um setor econômico. Neto, Cavalcanti et al. (2008) utilizaram índices setoriais do Petróleo e índices de desempenho médio do mercado brasileiro e alemão para a previsão da ação PETR4 da empresa Petrobrás.

3.3.2 Transformação das Variáveis

O processo de transformação consiste em aplicar operações matemáticas sobre o conjunto original dos dados. Esta transformação tem como objetivo eliminar informações redundantes ou irrelevantes, que não contribuem para o processo de classificação (WEBB, 2002). As transformações aplicadas sobre as variáveis, conforme observado em trabalhos correlacionados, são: taxa de variação (TAY; CAO, 2001) (MAGER; PAASCHE; SICK, 2008), médias móveis exponenciais (VIDOTTO; MIGLIATO; ZAMBON, 2009) (TAY; CAO, 2001) e médias móveis simples (BAHRAMY; CRONE, 2013) (NAEINI; TAREMIAN; HASHEMI, 2010).

A taxa de variação, um importante indicador de oscilação utilizado na análise técnica, representa o percentual de alteração que os valores sofrem em determinados períodos de tempo. Frequentemente a taxa de variação é escolhida como uma transformação a ser aplicada sobre os dados do modelo predictor. Tay e Cao (2001) descrevem essa alteração como uma vantagem importante, pois a distribuição dos dados transformados se tornará mais simétrica, se aproximando de uma distribuição normal, podendo possibilitar a obtenção de melhores resultados na previsão. Diferentemente da Equação 2.4, Mager, Paasche e Sick (2008) definiram a taxa de variação como

$$ROC_n = 100 \frac{v_t - v_{t-n}}{v_{t-n}} \quad (3.2)$$

onde ROC_n representa a variação do dia n em relação ao dia atual t , v_t é o valor da série no tempo t e v_{t-n} corresponde ao valor da série no tempo n anterior ao dia atual t . Mager, Paasche e Sick (2008) escolheram 5 taxas de variação para compor os dados de entrada do modelo predictor: ROC_1 , ROC_2 , ROC_3 , ROC_5 e ROC_8 . Já Tay e Cao (2001) escolheram 4 variações de preço para se integrarem ao conjunto de variáveis de entrada: ROC_5 , ROC_{10} ,

ROC_{15} e ROC_{20} . [Macedo et al. \(2013\)](#) empregaram o cálculo de variação de apenas 1 dia, ROC_1 , incrementando este dado aos outros dados já selecionados.

Médias móveis também são utilizadas para extrair características dos dados originais. [Naeini, Taremian e Hashemi \(2010\)](#) calcularam médias móveis simples para períodos de até 10 dias e utilizaram-nas como dados de entrada para o modelo de previsão. Já [Bahramy e Crone \(2013\)](#) empregaram a metodologia Bandas de *Bollinger*, que são formadas a partir do cálculo de médias móveis simples, para rastrear tendências e gerar sinais de compra/venda dos ativos. As informações obtidas através das Bandas de *Bollinger* utilizadas como entrada no modelo de previsão foram as séries que se referem as bandas central, inferior e superior da metodologia, assim como a série de sinais de compra e venda gerada.

Outro tipo de médias móveis, as exponenciais, também são utilizadas. [Tay e Cao \(2001\)](#) utilizaram o conceito de média móvel exponencial para manter o máximo de informação contido no preço de fechamento original quanto possível. Este processo foi necessário pois os outros dados de entrada do modelo são taxas de variação de preço, que podem remover algumas informações úteis dos valores originais. Por isso, calcularam a média móvel exponencial dos últimos 15 dias e descontaram seu valor do preço de fechamento do dia atual, acreditando que dessa forma este dado compensaria a perda de informações causadas pelo cálculo das taxas de variação. A escolha da média móvel exponencial é justificada pelo fato de ser melhor no acompanhamento de tendências se comparada à média móvel simples. Esse melhor desempenho ocorre porque a média móvel exponencial atribui maior peso aos dados mais recentes, reagindo com mais agilidade as mudanças do que a média móvel simples ([VIDOTTO; MIGLIATO; ZAMBON, 2009](#)).

Em alguns casos as variáveis selecionadas não passam por transformações. os dados são utilizados de forma bruta durante o processo de treinamento e previsão ([LI; XIONG, 2005](#)), ([MACIEL; BALLINI, 2009](#)), ([NETO; CAVALCANTI et al., 2008](#)), ([WU; LU, 2012](#)), ([KHAN; BANDOPADHYAYA; SHARMA, 2008](#)) e ([YILDIZ; YALAMA; COSKUN, 2008](#)).

3.4 PREVISÃO

Existem diversos métodos estatísticos que podem auxiliar na tarefa de previsão, como por exemplo: modelos auto-regressivos (AR), de médias móveis (MA) e modelos ARIMA ([RIBEIRO; GOLDSCHMIDT; CHOREN, 2009](#)). Além destes, técnicas de inteligência computacional tais como as redes neurais artificiais e Máquinas de Vetor de Suporte proporcionam a criação de metodologias avançadas de previsão. Em uma análise de trabalhos relacionados, nota-se que as escolhas dos modelos computacionais de previsão, em sua grande maioria, recaem sobre a técnica de Redes Neurais Artificiais.

3.4.1 Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) são um sistema de processamento de informações paralelo e distribuído, constituído de unidades de processamento simples, com objetivo de armazenar conhecimentos adquiridos através de experiências e torná-los disponíveis para uso (HAYKIN, 2001). As RNAs se assemelham ao cérebro humano pois os conhecimentos de ambos são adquiridos através de processos de aprendizagem. As conexões entre os neurônios, conhecidas como sinapses, tem associado a elas os pesos sinápticos que são os responsáveis por armazenar os conhecimentos adquiridos pela RNA.

As RNAs tem uma grande capacidade de generalização, produzindo saídas adequadas para dados de entrada que não estavam presentes durante o processo de aprendizagem (HAYKIN, 2001). Esta capacidade de generalização é essencial na previsão de séries temporais financeiras, uma vez que muito dificilmente os valores posteriormente fornecidos para a rede (modelo predictor) serão iguais aos valores que foram utilizados durante o processo de aprendizagem.

Luger (2013) afirma que RNAs são bem adequadas para tarefas de classificação, onde o objetivo é definir a qual classe pertence determinada entrada, tarefas de reconhecimento de padrões, identificando os padrões existentes nos dados e em tarefas de previsão, na geração de saídas a partir das entradas disponibilizadas. A partir disso pode-se concluir que o uso de redes neurais na previsão de tendência de séries financeiras é adequado, pois esta tarefa de previsão passa pelo reconhecimento dos padrões presentes nos dados de entrada e pela classificação destes dados quanto a tendência de movimentação dos preços.

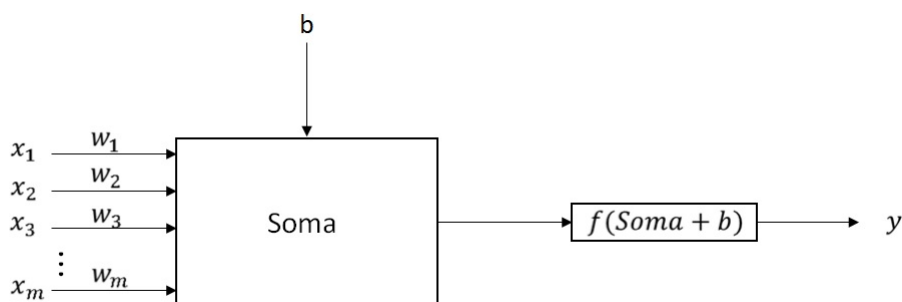
A base das RNAs é o neurônio artificial, que consiste na aplicação de uma função de ativação sobre a soma ponderada do neurônio, que é determinada pela força cumulativa dos sinais de entrada que chegam ao neurônio ponderados pelos pesos das conexões (LUGER, 2013). Um neurônio artificial pode ser definido como (HAYKIN, 2001)

$$y_k = \varphi(Soma_k + b_k) \quad (3.3)$$

$$Soma_k = \sum_{j=1}^m w_{kj}x_j \quad (3.4)$$

onde x_1, x_2, \dots, x_m são os sinais de entrada; w_1, w_2, \dots, w_m são os pesos sinápticos; m é o número de conexões que chegam ao neurônio k ; j é um neurônio da camada anterior; $Soma$ é a saída do somatório das entradas do neurônio ponderados pelos seus respectivos pesos sinápticos; b_k é o *bias*, valor aplicado externamente com o objetivo de alterar a entrada líquida da função de ativação; $\varphi(.)$ é a função de ativação que gera uma saída padronizada; e y_k é a saída do neurônio k . Uma representação gráfica de um neurônio artificial é exibida na [Figura 7](#).

Figura 7: Modelo genérico de um neurônio artificial.



Fonte: (LUGER, 2013) (HAYKIN, 2001) (Adaptado)

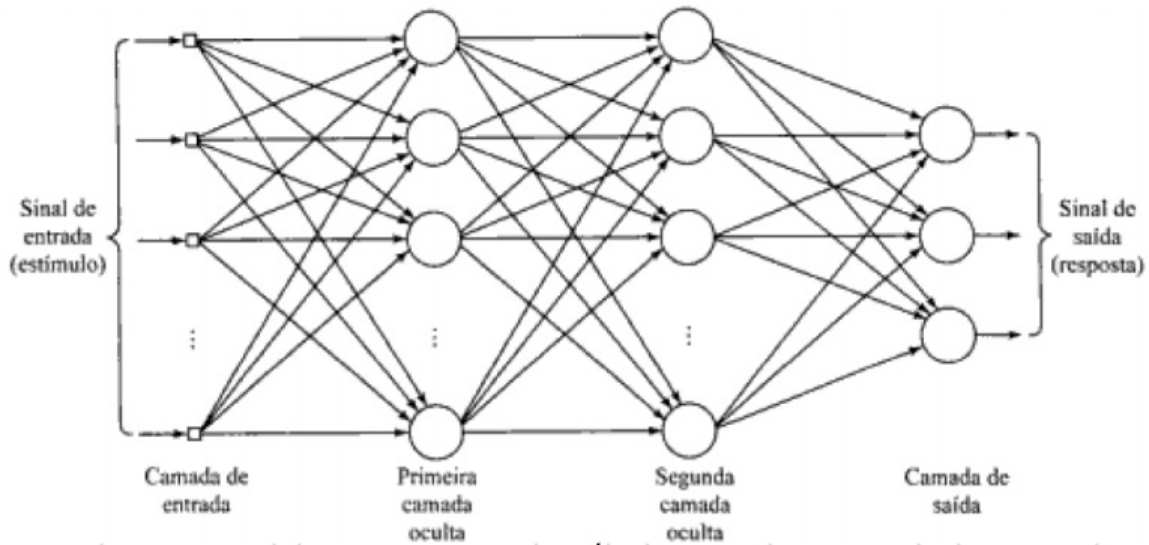
O *perceptron* é a forma mais simples de uma RNA e é capaz de resolver o problema da classificação de padrões linearmente separáveis¹. Consiste de um único neurônio com pesos sinápticos ajustáveis e um *bias*. Quando os problemas a serem solucionados são do tipo não linear, o uso de múltiplas camadas de *perceptrons* deve ser aplicado. Neste caso os neurônios de uma camada posterior, recebem a saída de todos os neurônios da camada anterior como suas entradas. Estas redes são conhecidas como RNA *Perceptrons* de Múltiplas Camadas (RNA-MLP).

As RNAs-MLP são as redes mais frequentemente modeladas na previsão de séries financeiras (VUI et al., 2013) (LI; MA, 2010) (YOO; KIM; JAN, 2005). A estrutura de uma RNA-MLP é composta por três tipos de camadas: entrada, oculta e saída. A camada de entrada é a responsável por receber os dados do ambiente externo e propagar para a próxima camada. Na sequência encontram-se as camadas ocultas, que não interagem diretamente com o ambiente externo. Normalmente, RNAs modeladas para previsão de séries financeiras utilizam uma (ABRAHAM; PHILIP; SARATCHANDRAN, 2004) (ZHANG; JIANG; LI, 2004) (FARIA et al., 2008) (KHAN; BANDOPADHYAYA; SHARMA, 2008) (NAEINI; TAREMIAN; HASHEMI, 2010) ou duas (MACEDO et al., 2013) (MACIEL; BALLINI, 2009) camadas ocultas. A última camada da rede é a camada de saída, que transfere os dados da RNA para o ambiente externo. A Figura 8 exemplifica uma arquitetura de RNA-MLP de 4 camadas, sendo duas delas ocultas.

Um dos maiores problemas de topologia das RNAs-MLP é a definição do número de neurônios integrantes de cada camada. Nas camadas de entrada e saída a definição é realizada de acordo com o problema ao qual a rede solucionará, o número de dados de entrada e de saída da RNA. O número de neurônios das camadas intermediárias, assim como o número de camadas ocultas, é definido de forma empírica, não havendo uma regra que defina a melhor arquitetura (BORGES, 2012). Na prática, a escolha do número de

¹ Padrões linearmente separáveis são aqueles em que os padrões se encontram em lados opostos de um hiperplano traçado.

Figura 8: Arquitetura de RNA-MLP com duas camadas ocultas.



Fonte: (HAYKIN, 2001, p. 186)

neurônios integrantes das camadas ocultas é determinada através da realização de testes com diferentes números de neurônios (NETO; CAVALCANTI et al., 2008) (FARIA et al., 2008). A arquitetura que apresenta o melhor desempenho nos testes é a escolhida.

3.4.1.1 Treinamento de Redes Neurais Artificiais

O processo de aprendizagem pode ser dividido em dois paradigmas: aprendizagem supervisionada e aprendizagem não supervisionada (HAYKIN, 2001). No processo de aprendizagem supervisionada há a presença de um "professor", alguém que possui conhecimentos sobre o assunto. O conhecimento é representado por um conjunto de exemplos de entrada-saída. Em virtude de seu conhecimento prévio, o "professor" é capaz de fornecer à rede neural uma resposta desejada, sendo possível determinar um sinal de erro entre a resposta obtida e a resposta desejada. Este processo de calcular o erro obtido e continuar ou não o processo de treinamento é conhecido como aprendizagem por correção de erro. No aprendizado não supervisionado a presença do "professor" inexistente, dessa forma os dados de entrada são colocados a disposição do modelo predictor para o seu treinamento, mas não se tem certeza se o modelo está sendo corretamente treinado.

Um dos algoritmos de aprendizagem supervisionada mais utilizados no treinamento de RNAs-MLP é o algoritmo de retropropagação, mais conhecido na literatura como algoritmo *backpropagation* (ZHANG; JIANG; LI, 2004) (MACEDO et al., 2013) (FARIA et al., 2008) (KHAN; BANDOPADHYAYA; SHARMA, 2008) (YILDIZ; YALAMA; COSKUN, 2008) (TAY; CAO, 2001). Este algoritmo é baseado na regra de aprendizagem por correção de erro e consiste na aplicação de dois passos através das diferentes camadas

da rede (HAYKIN, 2001): um passo para frente, a propagação, e um passo para trás, a retropropagação.

Na propagação, um vetor de entrada é aplicado aos neurônios da primeira camada. Todos os neurônios, exceto os da primeira camada, realizam o cálculo do seu valor de saída, conforme visto na Equação 3.3. A saída de cada neurônio é propagada para os neurônios da camada posterior, seguindo camada por camada, produzindo um conjunto de saídas da RNA. Durante o processo de propagação os pesos sinápticos, que são atribuídos de forma aleatória na inicialização da rede, não são alterados.

A retropropagação consiste em ajustar os pesos sinápticos, através de uma regra de correção de erro, de modo que a resposta obtida se mova para mais perto da resposta desejada (HAYKIN, 2001). O cálculo de erro de um neurônio j da camada de saída é definido por

$$e_j(n) = d_j(n) - o_j(n) \quad (3.5)$$

onde n se refere ao n -ésimo exemplo de treinamento, d_j se refere a resposta desejada, o_j corresponde ao valor obtido na saída do neurônio j e e_j é o erro na saída do neurônio j . Este sinal de erro é propagado para trás através da rede na direção contrária das conexões sinápticas, realizando o ajuste dos pesos sinápticos. A correção dos pesos sinápticos é definida como

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha[w_{ji}^{(l)}(n-1)] + \eta\delta_j^{(l)}(n)y_i^{(l-1)}(n) \quad (3.6)$$

onde $w_{ji}^{(l)}(n+1)$ é o novo peso da conexão sináptica entre o neurônio i da camada $l-1$ e o neurônio j da camada l para a próxima iteração $n+1$, $w_{ji}^{(l)}(n)$ é o peso atual da conexão, α é uma constante de momento, $w_{ji}^{(l)}(n-1)$ é o peso da conexão na iteração anterior $n-1$, η é a taxa de aprendizagem² definida, $y_i^{(l-1)}(n)$ é o valor de saída do neurônio i da camada anterior $l-1$ e $\delta_j^{(l)}$ é o gradiente local que é responsável por apontar as modificações necessárias nos pesos sinápticos, definido como

$$\delta_j^{(l)} = \begin{cases} e_j^{(L)}(n)\varphi'_j(v_j^{(L)}(n)) & \text{para o neurônio } j \text{ da camada de saída } L \\ \varphi'_j(v_j^{(l)}(n)) \sum_k \delta_k^{(l+1)}(n)w_{kj}^{(l+1)}(n) & \text{para o neurônio } j \text{ da camada oculta } l \end{cases} \quad (3.7)$$

onde e_j corresponde ao erro calculado na saída do neurônio j , L corresponde a camada de saída, l corresponde a uma camada oculta, n é a n -ésima iteração, $\varphi'_j(v_j^{(L)}(n))$ é a derivada da função de ativação φ do neurônio j , k é um neurônio de camada posterior a camada do neurônio j , w é o peso da conexão sináptica entre dois neurônios e $v_j(n)$ representa a soma

² Taxa de Aprendizagem é o valor que determina as variações dos pesos sinápticos a cada iteração. Quanto menor o valor da taxa de aprendizagem, mais suave serão as alterações nos pesos e mais demorado será o processo de aprendizagem. Se o valor for grande a rede aprenderá mais rápido mas pode se tornar instável.

ponderada de todas as entradas sinápticas acrescido do *bias* do neurônio j na iteração n , calculado por

$$v_j(n) = \sum_{i=0}^m w_{ji}(n)y_i n \quad (3.8)$$

onde m é o número total de entradas aplicadas ao neurônio j , o peso sináptico w_{j0} corresponde ao valor *bias* associado ao neurônio j , e y_i corresponde ao valor de saída do neurônio i .

O treinamento através do algoritmo *backpropagation* consiste em realizar o processo de propagação, seguido da retropropagação, para um conjunto de diferentes exemplos de treinamento, por diversas iterações, até que o critério de parada seja satisfeito (HAYKIN, 2001). O critério de parada pode ser alcançado através do uso de um conjunto de validação, que é um conjunto de dados formado por exemplos diferentes dos que são usados no treinamento. Estes exemplos são inseridos na rede e se o desempenho obtido com estes exemplos for bom o suficiente o treinamento é encerrado. Outra forma de definir o critério de parada é através da variação de erro entre as iterações, quando ela se tornar suficientemente pequena o treinamento é encerrado.

3.4.1.2 Função de Ativação

A função de ativação tem o propósito de restringir a amplitude de saída de um neurônio. Independente do valor v_j calculado no neurônio j a função de ativação irá transformá-lo em outro valor dentro de um intervalo específico, dependendo do tipo de função escolhida. Os tipos mais utilizados na previsão de séries financeiras são:

- Função de ativação sigmóide: é uma função não-linear muito utilizada em RNAs-MLP (ZHANG; JIANG; LI, 2004) (MACEDO et al., 2013) (KHAN; BANDOPADHYAYA; SHARMA, 2008) (TAY; CAO, 2001). A função sigmóide limita os dados a um intervalo $[0; 1]$. Esta limitação pode ser explicada através da plotagem de seu gráfico, conforme a Figura 9, onde percebe-se que ao aplicar a função sigmóide o valor de saída sempre está no intervalo $[0; 1]$, independente do valor de entrada. A função pode ser definida como

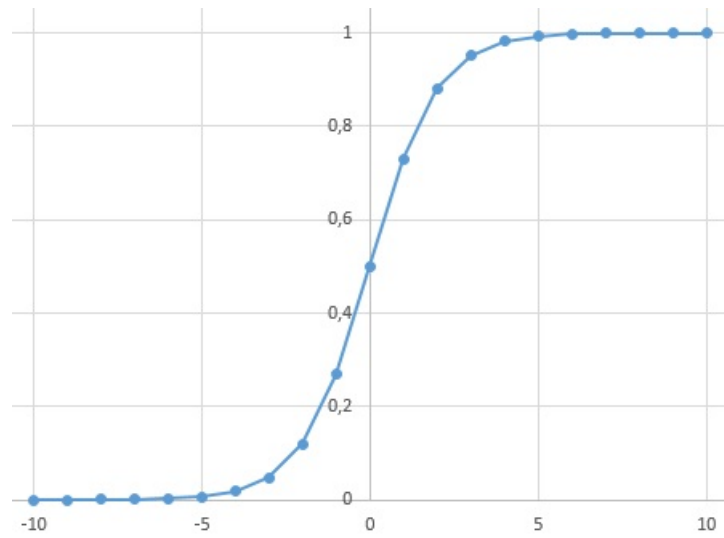
$$\varphi(v_j) = \frac{1}{1 + \exp^{-v_j}} \quad (3.9)$$

onde v_j é o valor da soma ponderada das entradas sinápticas do neurônio j .

- Função tangente hiperbólica: também bastante utilizada em RNAs-MLP por sua não-linearidade (FARIA et al., 2008) (NETO; CAVALCANTI et al., 2008), limita os valores a um intervalo $[-1; 1]$. A Figura 10 exhibe o gráfico da função tangente hiperbólica. Pode ser definida como

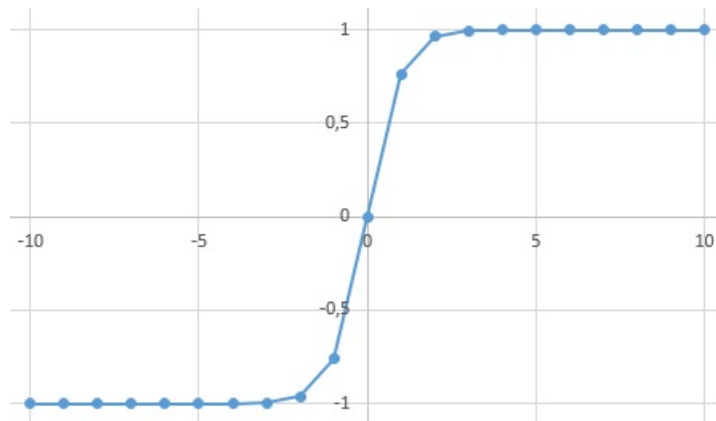
$$\varphi(v_j) = \tanh(v_j) \quad (3.10)$$

Figura 9: Gráfico da Função Sigmóide.



Fonte: desenvolvido pelo autor

Figura 10: Gráfico da Função Tangente Hiperbólica.



Fonte: desenvolvido pelo autor

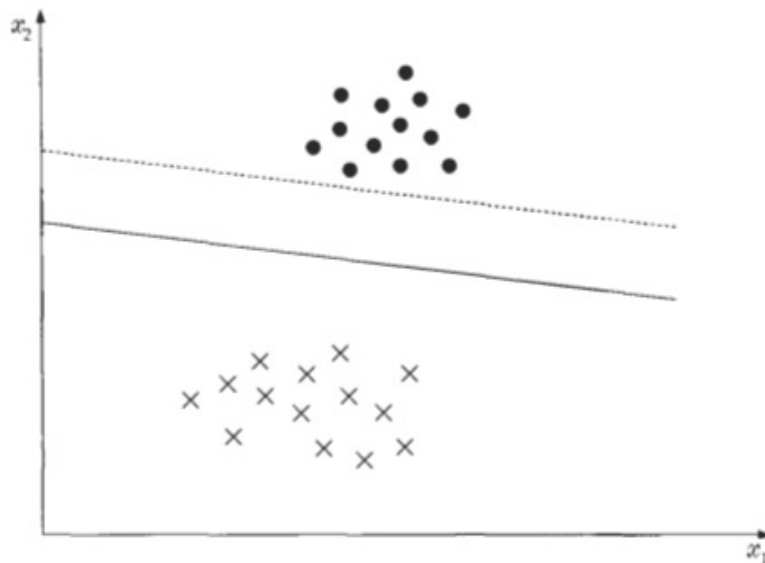
onde v_j é o valor da soma ponderada das entradas sinápticas do neurônio j .

3.4.2 Máquinas de Vetor de Suporte

As máquinas de vetor de suporte (SVM), assim como as RNAs-MLP, são máquinas de aprendizagem que podem ser usadas na classificação de padrões e na regressão linear (HAYKIN, 2001). A SVM está fundamentada na teoria da aprendizagem estatística, sendo uma implementação do método de minimização estrutural de risco. Este método de minimização estrutural de risco é baseado no fato de que a taxa de erro de uma máquina de aprendizagem, aplicada sobre dados de teste, é limitada pela soma da taxa de erro da máquina durante o treinamento e um termo que depende da *dimensão de Vapnik-Chervonenkis* (V-C) (HAYKIN, 2001).

Para explicar seu funcionamento, pega-se, como exemplo, a classificação de padrões linearmente separáveis, entre duas possíveis classes. O objetivo dos classificadores é a concepção de um hiperplano, que normalmente não é o único existente, que classifique corretamente todos os vetores de treinamento como mostra a Figura 11 (THEODORIDIS; KOUTROUMBAS, 2003). Nota-se nesta figura que existem 2 hiperplanos traçados que separam os dados de classes diferentes. Porém o hiperplano, representado pela linha contínua, possui uma maior margem de erro, de modo que os dados de ambas as classes podem mover-se um pouco mais livremente, com menos risco de causar um erro. Em outras palavras, o classificador da linha contínua tem uma maior capacidade de generalização, o que permite ele operar satisfatoriamente com dados diferentes dos dados de treinamento (THEODORIDIS; KOUTROUMBAS, 2003).

Figura 11: Um exemplo de um problema de 2 classes linearmente separáveis com dois classificadores lineares possíveis.



Este é o objetivo principal das SVMs, a construção de um hiperplano de modo que a separação entre os padrões diferentes seja máxima (HAYKIN, 2001). Sob esta condição, este hiperplano é chamado de hiperplano ótimo, que é único, e possui a maior capacidade de generalização na classificação de padrões.

As amostras de treinamento apresentadas às SVMs se assemelham àquelas apresentadas às RNAs, pois são do tipo (x, d) , onde x é o valor de entrada e d representa a resposta desejada. Assumindo que os padrões existentes são $d = -1$ e $d = 1$, pode-se representar a equação do hiperplano como (HAYKIN, 2001)

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (3.11)$$

onde \mathbf{w} é um vetor de peso ajustável, b é um *bias*. O treinamento consiste em ajustar os pesos do vetor de forma que a separação dos padrões seja ótima. Pode-se representar, de maneira simplificada, a classificação de padrões da seguinte forma

$$\begin{aligned} \mathbf{w}^T \mathbf{x} + b &\geq +1 \text{ para } d = +1 \\ \mathbf{w}^T \mathbf{x} + b &\leq -1 \text{ para } d = -1 \end{aligned} \quad (3.12)$$

Os pontos de dados (x, d) para os quais a primeira ou a segunda linha da Equação 3.12 é satisfeita com o sinal de igualdade são chamados de *vetores de suporte* (HAYKIN, 2001). Estes vetores são aqueles pontos de dados que se encontram mais próximos da superfície de decisão e são, portanto, os mais difíceis de classificar.

A classificação de tendências em séries temporais financeiras não se encaixa na modalidade de reconhecimento de padrões linearmente separáveis. Porém, as SVMs também são capazes de encontrar hiperplanos ótimos para padrões não-linearmente separáveis, nos quais se busca minimizar o erro de classificação (HAYKIN, 2001). Neste caso podem ser utilizados os métodos multiplicadores de Lagrange para encontrar o vetor de pesos \mathbf{w} e o *bias* que minimizam o erro de classificação e/ou o teorema de Cover que transforma o espaço multidimensional em um novo espaço de características onde os padrões são linearmente separáveis. Como esta parte é mais complexa não será aprofundada neste trabalho.

Esta técnica de inteligência computacional foi utilizada por Tay e Cao (2001) na previsão dos valores de índices futuros de ações e de títulos públicos federais dos Estados Unidos da América e da Alemanha, com vencimentos em 10 e 30 anos. Também foi empregada em alguns outros trabalhos que tinham por objetivo prever outras séries temporais financeiras (MAGER; PAASCHE; SICK, 2008) (BAHRAMY; CRONE, 2013).

3.4.3 Outros Modelos Preditores

Além das técnicas já citadas, existem outras que podem ser utilizadas na modelagem de sistemas previsores de séries temporais financeiras. Dentre estas pode-se destacar o uso da Lógica *Fuzzy* e dos Algoritmos Genéticos.

A Lógica *Fuzzy*, ou lógica nebulosa, tem como objetivos modelar modos de raciocínio aproximados ao invés de precisos (MARQUES; GOMES, 2009). Seu funcionamento é baseado em regras, definidas com palavras e não com números. A partir da aplicação destas regras é que se define o quão membro determinada situação é de um determinado conjunto (CAMARGOS, 2002). A faixa de transição entre membro e não membro de um conjunto é representada por um intervalo numérico, que vai de "0" (totalmente não membro) a "1" (totalmente membro). A maior dificuldade desta técnica é definir os termos linguísticos e regras que controlam o conjunto. Sistemas desenvolvidos com lógica *fuzzy* geralmente são híbridos, onde uma outra técnica é utilizada para formar o sistema, como redes neurais artificiais (LI; XIONG, 2005) ou algoritmos genéticos (MARQUES; GOMES, 2009).

Algoritmos genéticos são baseados no processo de seleção natural (evolução biológica) (KHAN; BANDOPADHYAYA; SHARMA, 2008). Para solucionar um problema com algoritmos genéticos, são geradas aleatoriamente populações iniciais, que são medidas através de uma função de avaliação. Depois são aplicados mecanismos de reprodução (mutação e cruzamento) sobre as melhores populações, gerando novas populações, enquanto as piores são descartadas. Este ciclo de reprodução, avaliação e descarte é repetido até encontrar uma população que solucione o problema (WUERGES; BORBA, 2010). Os algoritmos genéticos também são utilizados associados a outras técnicas como redes neurais artificiais, nas quais podem auxiliar no ajuste dos pesos sinápticos (KHAN; BANDOPADHYAYA; SHARMA, 2008).

3.5 AVALIAÇÃO DOS RESULTADOS

Após a definição do modelo de previsão são definidas métricas para checar o desempenho da previsão, analisando o quão bom ou ruim está a previsão através do modelo desenvolvido (MORETTIN; TOLOI, 1987). As métricas de análise de desempenho frequentemente utilizadas na avaliação dos resultados de sistemas previsores podem ser divididas em dois grupos: métricas de análise do valor nominal do ativo e métricas de análise do acerto da tendência de movimentação dos preços.

As medidas de desempenho são expressas através do erro de previsão. O erro de previsão consiste na diferença entre o valor desejado (*alvo*) e o valor obtido na saída do previsor (*obtido*). Dessa forma podemos definir o erro de previsão pela equação

$$e_t = (alvo_t - obtido_t) \quad (3.13)$$

onde $alvo_t$ é o valor desejado como saída do preditor no tempo t , e $obtido_t$ é o valor obtido na saída do preditor em um tempo determinado t .

3.5.1 Avaliação do Desempenho na Previsão de Valor Nominal

Na avaliação do desempenho dos previsores de valor nominal de ativos financeiros, algumas métricas são frequentemente utilizadas para comparar o desempenho entre os modelos previsores desenvolvidos. Essas métricas estão descritas a seguir.

Erro Quadrático Médio

O Erro Quadrático Médio, do Inglês *Mean Square Error* - MSE, é uma medida de comparação entre dois valores, ou duas séries temporais, que fornece uma pontuação quantitativa descrevendo o grau de semelhança, ou de distorção, entre os valores (WANG; BOVIK, 2009). Esta métrica foi empregada para analisar o desempenho dos modelos elaborados por Neto, Cavalcanti et al. (2008) e Naeini, Taremian e Hashemi (2010).

A equação para obter o MSE pode ser definida como (WANG; BOVIK, 2009)

$$MSE(\text{obtido}, \text{alvo}) = \frac{1}{n} \sum_{t=1}^n (\text{obtido}_t - \text{alvo}_t)^2 \quad (3.14)$$

onde n é a quantidade de pontos da série temporal. Quanto menor for o resultado do MSE, melhor é o modelo de previsão. MSE igual a zero equivale a um modelo que prevê o resultado com exatidão.

Erro Quadrático Médio Normalizado

O Erro Quadrático Médio Normalizado, do Inglês *Normalized Mean Square Error* - NMSE, mede a relação entre o quadrado do erro de previsão do sistema testado e o quadrado dos erros de um modelo do tipo *Random Walk* (Caminho Aleatório) (NETO; CAVALCANTI et al., 2008). Pelo modelo *Random Walk* a previsão de um determinado valor da série no instante t é calculada a partir de uma dependência linear com o seu último valor no tempo $t-1$. Dessa forma chega-se a $\text{obtido}_t = \text{obtido}_{t-1} - r_t$, onde r_t é um valor aleatório gerado por uma distribuição uniforme de números aleatórios. O NMSE pode ser definido como (NETO; CAVALCANTI et al., 2008)

$$NMSE = \frac{\sum_{t=1}^n (\text{alvo}_t - \text{obtido}_t)^2}{\sum_{t=1}^n (\text{obtido}_t - \text{obtido}_{t-1})^2}. \quad (3.15)$$

Quando $NMSE \geq 1$, o modelo testado é ineficiente pois tem desempenho igual ou inferior ao *Random Walk*. Se o $NMSE < 1$ o modelo testado é mais eficiente que o modelo *Random Walk*. Quanto mais próximo de zero o valor de NMSE, melhor é o desempenho do sistema predictor. Esta métrica foi adotada por Tay e Cao (2001) e Neto, Cavalcanti et al. (2008) em seus trabalhos.

Raiz do Erro Quadrático Médio

A métrica Raiz do Erro Quadrático Médio, do Inglês *Root Mean Square Error* -

RMSE, também é utilizada para verificar a eficiência das previsões servindo de comparação de desempenho entre os modelos desenvolvidos (BAHRAMY; CRONE, 2013). RMSE pode ser definido como (FARIA et al., 2008)

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (alvo_t - obtido_t)^2}{n}}. \quad (3.16)$$

Além de FARIA et al. (2008) e Bahramy e Crone (2013), Wu e Lu (2012) e Naeini, Taremian e Hashemi (2010) também utilizaram esta métrica para avaliar o desempenho de seus modelos de previsão. Quanto menor o valor de RMSE melhor é o modelo de previsão.

Erro Médio Percentual Absoluto

O Erro Médio Percentual Absoluto, do Inglês *Mean Absolute Percentage Error* - MAPE, apresenta o erro de previsão do modelo em porcentagem. MAPE pode ser definido como

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left[\frac{alvo_t - obtido_t}{alvo_t} \right] * 100. \quad (3.17)$$

Apresentar o resultado em percentual é o grande diferencial do MAPE em relação ao MSE, uma vez que na métrica MSE o resultado deve ser interpretado, levando-se em conta a escala dos valores avaliados. Já na métrica MAPE, o resultado é facilmente entendido, pois representa o percentual de erro do modelo. Quanto menor o MAPE, melhor é o desempenho do previsor.

A métrica MAPE é frequentemente utilizada na avaliação de sistemas previso- res (ABRAHAM; PHILIP; SARATCHANDRAN, 2004) (NETO; CAVALCANTI et al., 2008) (WU; LU, 2012) (NAEINI; TAREMIAN; HASHEMI, 2010).

3.5.2 Avaliação do Desempenho na Previsão de Tendência

A avaliação dos previso- res de tendência pode ser encarada como uma avaliação de classificadores. As possíveis classes que uma situação pode pertencer são: tendência de alta ou tendência de baixa dos preços.

Taxa de Acerto

A Taxa de Acerto, do Inglês *Hit Rate*, corresponde a uma métrica de análise de desempenho utilizada quando o universo de possibilidades (classes) é definido. A previsão de tendência do preço de séries financeiras se encaixa neste padrão, pois ou o preço cai ou o preço sobe, muito raramente o preço se mantém estável. Por ser uma previsão de valores financeiros, se o preço se mantém estável, pode-se considerar que o modelo acertou a previsão, pois a estabilidade não afeta o objetivo final que é a obtenção de retorno financeiro.

A equação para medir a taxa de acerto varia de acordo com os valores tratados no modelo predictor. Em [Li e Xiong \(2005\)](#) o sistema de previsão retorna um valor que corresponde ao preço futuro da série financeira analisada. A equação para avaliar o percentual de acerto da direção foi assim definida

$$HIT = \frac{1}{N} \sum_{t=1}^n D_t \quad (3.18)$$

onde

$$D_t = \begin{cases} 1, & \text{se } (alvo_t - alvo_{t-1}) * (obtido_t - obtido_{t-1}) \geq 0 \\ 0, & \text{caso contrário} \end{cases} . \quad (3.19)$$

No modelo proposto por [Yildiz, Yalama e Coskun \(2008\)](#) o cálculo da direção D_t teve de ser alterado, pois o sistema predictor retorna um valor no intervalo $[0; 1]$. A [Equação 3.18](#) pode ser mantida, sem alterações, mas a [Equação 3.19](#), de verificação de acerto de tendência, deve ser alterada como

$$D_t = \begin{cases} 1, & \text{se } obtido_t > 0.5 \\ 0, & \text{caso contrário} \end{cases} . \quad (3.20)$$

Nem todos os autores utilizam a mesma nomenclatura quando se referem a métrica Taxa de Acerto. [Li e Xiong \(2005\)](#) utilizaram o nome *Ratio of the Correct Tendency*, na tradução livre Razão da Tendência Correta, mas o resultado da aplicação desta métrica também representa o percentual de acerto na previsão de tendência. *Prediction On Change In Direction*, que pode ser traduzido por Previsão de Alteração da Direção, é outra nomenclatura para o cálculo de Taxa de Acerto, utilizada por [Neto, Cavalcanti et al. \(2008\)](#). Além dos trabalhos já citados, [Khan, Bandopadhyaya e Sharma \(2008\)](#) e [Mager, Paasche e Sick \(2008\)](#) também propuseram métricas similares com o mesmo objetivo.

Apesar de terem nomes diferentes, essas métricas representam o percentual de acerto da mudança de direção dos preços. Quanto maior o percentual de acerto melhor é o desempenho do predictor. Pode-se considerar que se o percentual de acerto for inferior a 50% o modelo proposto não deve ser utilizado, pois uma escolha aleatória sem nenhum critério poderia ter desempenho melhor que o sistema predictor.

Soma das Perdas e Ganhos

A Soma de Perdas e Ganhos, do Inglês *Sum of Losses and Gains* - SLG, foi proposta por [Neto, Cavalcanti et al. \(2008\)](#) e representa uma boa forma de medir o desempenho quando o assunto é previsão de séries financeiras. Quando a previsão acerta a tendência de movimento do preço, o valor absoluto $|alvo_t - alvo_{t-1}|$ é somado, caso contrário $|alvo_t - alvo_{t-1}|$ é subtraído. O SLG é calculado por

$$SLG = \frac{1}{n} \sum_{t=1}^n L_t \quad (3.21)$$

onde L_t é calculado por

$$L_t = \begin{cases} + |alvo_t - alvo_{t-1}|, & \text{se } (alvo_t - alvo_{t-1}) * (obtido_t - obtido_{t-1}) > 0 \\ - |alvo_t - alvo_{t-1}|, & \text{caso contrário} \end{cases} \quad (3.22)$$

O SLG representa a obtenção de lucro ou prejuízo. Se o modelo proposto tem um desempenho com $SLG \leq 0$ ele não deve ser utilizado pois representaria uma perda de capital. Quanto maior o SLG melhor é o desempenho do modelo.

3.6 CONSIDERAÇÕES FINAIS

Neste capítulo foram abordados assuntos referentes à previsão de séries temporais financeiras. Através da análise de trabalhos correlacionados pode-se determinar os aspectos mais importantes na definição de modelos previsores, como a seleção das variáveis, a transformação sobre elas aplicada, a modelagem de redes neurais e as formas de avaliação do desempenho dos previsores.

A escolha correta das variáveis é fundamental para obter, ao final, um bom sistema predictor. As variáveis devem carregar consigo informações que permitam, com o auxílio de um bom reconhecedor de padrões, classificar os dados conforme padrões pré-definidos. O processo de transformação das variáveis também pode auxiliar nessa tarefa. Através da aplicação de operações matemáticas e indicadores de análise técnica é possível agregar mais informações de classificação aos dados.

A técnica computacional mais utilizada na previsão de séries temporais financeiras é a RNA-MLP. Seu treinamento normalmente é realizado através do algoritmo *backpropagation*. Os principais problemas encontrados na criação de RNAs são as definições de sua topologia. Normalmente o número de camadas é definido entre três e quatro. O número de neurônios de cada camada escondida é definido de forma empírica, inicialmente são testadas diversas configurações partindo de uma quantidade pequena de neurônios (menor que o número de classes existentes) até uma quantidade grande (maior que o dobro do número de dados de entrada). Posteriormente são analisados os melhores resultados e é definido um intervalo do número de neurônios a ser testado em cada experimento.

Por fim, neste capítulo, foram apresentadas as principais métricas de análise de desempenho de sistemas previsores de séries temporais financeiras. Elas podem ser divididas em dois grupos, métricas de análise de previsores de valor nominal e métricas de análise dos previsores de tendência.

4 PROPOSTA DE SOLUÇÃO

Neste capítulo é apresentada a proposta de solução para o problema de previsão de séries temporais financeiras. Na [seção 4.1](#) é apresentada a arquitetura proposta para o previsor de tendência de séries temporais financeiras. A [seção 4.2](#) define as séries temporais financeiras escolhidas para serem analisadas neste trabalho. A [seção 4.3](#) especifica quais são as variáveis utilizadas como fonte de informações para a previsão. Por fim, a [seção 4.4](#) relaciona as métricas que serão utilizadas para analisar o desempenho do sistema previsor proposto.

4.1 ARQUITETURA DO SISTEMA

Este trabalho tem como objetivo principal fazer a combinação de sistemas especialistas para prever o comportamento futuro de séries temporais financeiras. Os sistemas especialistas se caracterizam por adquirir conhecimentos, através de um processo de aprendizagem, para utilizá-los posteriormente na solução de problemas similares.

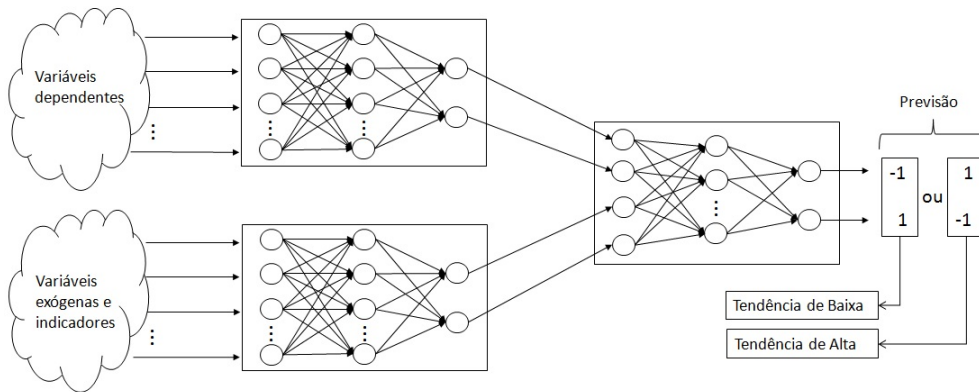
A previsão de séries temporais financeiras pode ser considerada uma tarefa complexa. A resolução de uma tarefa complexa, segundo [Haykin \(2001\)](#), pode ser obtida através da divisão em sub-tarefas. Estas podem ser delegadas à sistemas especialistas. Os resultados obtidos por cada especialista são combinados gerando uma solução supostamente melhor do que qualquer solução alcançável por um especialista atuando isoladamente.

Dentre as técnicas de inteligência computacional utilizadas para obtenção de um sistema especialista na previsão de séries temporais financeiras, a que mais se destaca são as RNAs. Esse destaque se dá pelo fato de que as RNAs possuem uma grande capacidade de generalização, sendo eficientes na previsão de séries temporais ([RIBEIRO; GOLDSCHMIDT; CHOREN, 2009](#)).

Neste contexto, é proposto um modelo de combinação de sistemas especialistas como mostra a [Figura 12](#). Este sistema foi modelado através de três RNAs-MLP, objetivando prever o comportamento futuro de séries temporais financeiras. A primeira RNA-MLP utiliza, como dados de entrada, as variáveis dependentes do ativo. A segunda RNA utiliza dados exógenos, que são independentes do ativo mas que podem influenciar na formação do preço do ativo, e dados gerados por indicadores de análise técnica. A terceira RNA, é uma RNA de fusão, que utiliza os dados de saída da primeira e segunda RNA como seus dados de entrada. Dessa forma, as duas primeiras RNAs são treinadas para fazer a previsão individualmente. Seus resultados são combinados em uma RNA de fusão, com o objetivo de melhorar a previsão, agregando o conhecimento de ambas as RNAs anteriores.

As três RNAs fazem a previsão da tendência de movimentação do preço. Quando os

Figura 12: Arquitetura do Sistema Previsor Proposto.



Fonte: desenvolvido pelo autor

valores de saída dos neurônios superior e inferior de cada RNA forem -1 e 1, respectivamente, representarão uma tendência de baixa do preço do ativo. Quando a saída for 1 e -1 para os neurônios superior e inferior, respectivamente, a tendência é de alta no preço do ativo.

4.2 SELEÇÃO DAS SÉRIES FINANCEIRAS

Séries temporais financeiras são um conjunto de dados históricos, que indicam a cotação dos mais diversos ativos no mercado financeiro ao longo do tempo, como série de cotação de moedas estrangeiras, de índices de mercado, de produtos agropecuários e de outros ativos. Neste trabalho foram selecionadas séries temporais financeiras de ações do mercado brasileiro para serem analisadas. A seleção das ações utilizadas neste trabalho seguiu os critérios:

- As ações devem apresentar alta liquidez no mercado acionário: segundo [CORREIA, AMARAL e BRESSAN \(2008\)](#), o nível de liquidez de um ativo reflete a facilidade com que ele é negociado, sem prejuízo de valor. É por este motivo, que são escolhidas as ações de maior liquidez.
- As ações selecionadas devem pertencer a empresas e setores de atuação diferenciados: diferentes segmentos do mercado. É utilizada a setorização da BM&FBovespa para fins de comparação.
- As ações devem estar sendo negociadas na BM&FBovespa desde 01 de janeiro de 2011.

Como forma de avaliação do nível de liquidez das ações, foi considerado o volume financeiro negociado da ação nos últimos 12 meses. A [Tabela 1](#) exibe o *ranking* das 10

ações com maior participação em volume financeiro negociado dentre as participantes da carteira do IBovespa.

Tabela 1: *Ranking* das 10 Ações com Maior Percentual de Participação em Volume Financeiro da Carteira de Ações do IBovespa.

Código ▾	Ação ▾	Tipo ▾	Part. (%) (2) ▾
ITUB4	ITAUUNIBANCO	PN ED N1	9,980
PETR4	PETROBRAS	PN	8,700
BBDC4	BRADESCO	PN N1	7,640
ABEV3	AMBEV S/A	ON	6,640
PETR3	PETROBRAS	ON	5,534
VALE5	VALE	PNA N1	4,503
BRFS3	BRF SA	ON NM	3,509
VALE3	VALE	ON N1	3,394
ITSA4	ITAUSA	PN ED N1	3,130
BBAS3	BRASIL	ON NM	2,676

Fonte: (BM&FBOVESPA, 2014)

Seguindo os critérios definidos e, limitando o número de ações a serem analisadas em três, as selecionadas foram:

- ITUB4: ação do tipo preferencial da empresa Itaú Unibanco. A Itaú Unibanco é uma empresa que atua no setor financeiro.
- PETR4: ação do tipo preferencial da empresa Petrobrás. A Petrobrás é uma empresa de capital aberto estatal, pois o acionista majoritário é o Governo Federal. Atua no segmento de energia, nas áreas de exploração, produção, refino, comercialização e transporte de petróleo, gás natural e seus derivados.
- VALE5: ação do tipo preferencial da empresa Vale S.A.. A Vale é a maior produtora de minério de ferro do mundo e atua também em outros segmentos minerais.

A ação BBDC4 da empresa Bradesco não foi selecionada por pertencer ao mesmo setor de atuação da ITUB4, o setor financeiro. A ação ABEV3, a 4^a no *ranking* das ações de maior liquidez, não foi selecionada pois passou a ser negociada na BM&FBovespa apenas em novembro de 2013, e neste trabalho serão considerados os dados de negociação desde 01/01/2011. A 5^a ação do *ranking*, a PETR3, também pertence a Petrobrás. Dessa forma, a ação VALE5 foi selecionada pois é a 3^a de maior liquidez que atende aos critérios de escolha estabelecidos.

4.3 SELEÇÃO DAS VARIÁVEIS DE ENTRADA

A seleção das variáveis é parte fundamental no projeto de um sistema especialista (FUKUNAGA, 1990). É das variáveis de entrada que são extraídas as características inerentes ao comportamento dos preços das ações. As características presentes nas variáveis, aliadas a um bom sistema classificador, permitem o reconhecimento dos padrões existentes nos dados. A partir deste reconhecimento é possível prever valores futuros. Neste trabalho as variáveis selecionadas para a previsão são separadas em três grupos: variáveis dependentes do ativo, indicadores de análise técnica e variáveis exógenas.

As variáveis dependentes do ativo a ser previsto representam um resumo diário da negociação do ativo. São fornecidas pelas bolsas de valores onde os ativos são negociados. A seguir estão relacionadas as variáveis dependentes utilizadas como entrada no sistema previsor:

- Cotação de fechamento diário da ação;
- Preço de abertura no dia;
- Preço máximo negociado no dia;
- Preço mínimo negociado no dia;
- Volume negociado no dia.

Sinais de compra e venda de ações, gerados a partir de indicadores de análise técnica, podem ser utilizados como em Bahramy e Crone (2013). Porém, os sinais de compra e venda, que são gerados através do cruzamento de médias móveis, não ocorrem diariamente. A presença de variáveis de entrada redundantes e/ou constantes prejudica o desempenho da rede neural, dificultando o processo de aprendizagem e tornando o treinamento mais difícil e demorado (LIMA, 2015). Por este motivo, ao invés de utilizar os sinais de compra e venda gerados, optou-se por utilizar os próprios indicadores responsáveis pela geração destes sinais. Os indicadores selecionados foram:

- Bandas de Bollinger;
- MACD;
- Intersecção de Médias Móveis.

Os dados exógenos frequentemente selecionados como dados de entrada para modelos previsores são taxas de câmbio de moedas e índices de desempenho médio fornecido pelas bolsas de valores. Neste contexto e considerando as ações selecionadas os dados exógenos escolhidos foram:

- Índice IBovespa;
- Índices Setoriais: para as ações ITUB4 e VALE5 existem índices de setor fornecidos pela BM&FBovespa, o índice financeiro (IFNC) e o índice de materiais básicos (IMAT), respectivamente. No caso da PETR4, a BM&FBovespa não possui um índice de setor. Por isso foi selecionado, assim como em [Neto, Cavalcanti et al. \(2008\)](#), a cotação do barril de petróleo *Brent* em substituição ao índice setorial, pois é uma importante informação neste setor.
- Cotação do Dólar: selecionada pois o valor da moeda americana influencia nos preços de muitos produtos no mercado brasileiro. Desta forma, esta pode ser uma importante informação na precificação de algumas ações.

As variáveis selecionadas são divididas entre as RNAs-MLP, conforme o modelo exposto na [Figura 12](#). A primeira RNA-MLP utiliza como dados de entrada as variáveis dependentes do ativo. Os indicadores de análise técnica, bem como as variáveis exógenas selecionadas, são os dados de entrada da segunda RNA-MLP.

4.4 AVALIAÇÃO DO DESEMPENHO

A partir do conjunto de dados de teste é realizada uma avaliação do desempenho do previsor. As métricas de avaliação consideram o acerto ou erro da previsão de movimentação do preço. Dentre as métricas expostas na [seção 3.5](#), foi selecionada a métrica taxa de acerto. Já a métrica soma de perdas e ganhos, foi adaptada para apresentar valores que representem melhor as perdas e ganhos.

A métrica taxa de acerto representa o percentual de acerto da previsão da tendência dos preços, sendo uma medida que representa a confiabilidade do sistema. Considera-se aceitável o desempenho de um previsor com taxa de acerto superior a 50%, já que a probabilidade de acertar é 50%. Quanto maior o percentual de acerto obtido mais promissor é o sistema previsor.

Como neste trabalho o previsor gera a tendência do preço e não o preço nominal, o cálculo foi adaptado conforme

$$HIT = \frac{1}{n} \sum_{t=1}^n D_t \quad (4.1)$$

$$D_t = \begin{cases} 1, & \text{se } ((valor_t - valor_{t-1}) \geq 0 \text{ e } [obtido_1, obtido_2] = [1, -1]) \\ 1, & \text{se } ((valor_t - valor_{t-1}) < 0 \text{ e } [obtido_1, obtido_2] = [-1, 1]) \\ 0, & \text{caso contrário} \end{cases} \quad (4.2)$$

onde HIT corresponde à taxa de acerto, n equivale ao número de testes realizados, D_t corresponde ao número de acertos na previsão da tendência, $valor_t$ é o preço do ativo no dia t , $obtido_1$ e $obtido_2$ são os valores de saída dos neurônios da camada de saída da RNA.

Por sua vez, a métrica soma de perdas e ganhos, se apresenta como uma possibilidade de avaliar o desempenho de previsores de tendência de movimentação dos preços de ativos financeiros. No entanto, essa métrica não considera uma variável importante no cálculo das perdas e ganhos no mercado à vista: a posse de ações. Pela métrica SLG, as perdas e ganhos são somadas conforme o acerto de tendência, sem considerar se o investidor tem ações para vender ou capital para comprar. Por exemplo, se o sistema predictor indica tendência de baixa em dias consecutivos, o investidor só terá um ganho no primeiro dia, pois nos dias seguintes ele já não possuirá mais ações para vender. A métrica SLG desconsidera a posse de ações para calcular o ganho. Portanto, visando ter um valor mais próximo do real, foi proposta uma simulação de investimento, partindo de um capital inicial a ser investido e comparando, ao final do período analisado, com o capital obtido. A simulação segue as regras:

- Comprar o maior número de ações possível de acordo com o capital disponível, caso o sistema predictor indique uma tendência de alta;
- Vender todas as ações, de posse do investidor, caso o sistema indique uma tendência de baixa.

Para que a simulação de investimento seja possível, foi necessário fixar uma regra quanto ao preço de compra e venda das ações. O preço praticado é igual ao preço de fechamento do dia atual. Na prática isso não é possível, pois após o fechamento do pregão só é possível negociar ações no dia seguinte, e o preço de abertura do dia seguinte é diferente do preço de fechamento do dia atual. Por isso, considera-se que a previsão de tendência ocorre minutos antes do fechamento do pregão, de forma que os preços praticados naquele instante não diferem muito dos preços de fechamento e que a negociação (compra ou venda) ocorra ainda no mesmo dia.

Ao final do período analisado, é calculado o percentual de retorno com a simulação de investimento. Um percentual de 100% significa que o capital inicial foi mantido ao final do período. Percentuais maiores que 100% representam lucro e percentuais menores que 100% representam prejuízo com o investimento durante o período analisado.

5 DESENVOLVIMENTO E RESULTADOS

Neste capítulo são apresentadas as etapas do desenvolvimento da proposta de solução para o problema de previsão de séries temporais financeiras e os resultados alcançados. A [seção 5.1](#) trata sobre a coleta dos dados e sobre os pré-processamentos realizados. Na [seção 5.2](#) o assunto abordado é a configuração das RNAs, tratando de aspectos como arquitetura, parâmetros da rede, critérios de parada do treinamento, função de ativação e a necessidade de normalização dos dados de entrada. A [seção 5.3](#) apresenta os experimentos realizados na busca por um sistema previsor eficiente para a RNA com variáveis dependentes e para a RNA com variáveis exógenas e indicadores. Por fim, na [seção 5.4](#) são relatados os experimentos realizados através da fusão de características e dos sistemas especialistas obtidos na seção anterior.

5.1 COLETA E PRÉ-PROCESSAMENTO DOS DADOS

Como a análise técnica está fundamentada na hipótese que preços passados influenciam nos preços futuros, é necessária a coleta de dados históricos referente a negociação dos ativos. Este trabalho considerou os dados históricos de 01 de janeiro de 2011 a 10 de março de 2014, gerando 750 amostras, para o treinamento e teste do modelo previsor. Nos trabalhos correlatos o período escolhido é muito variável, em alguns casos foram utilizadas 2000 amostras (aproximadamente 8 anos) enquanto em outros a previsão foi realizada baseada em 250 amostras (aproximadamente 1 ano).

Os dados coletados foram armazenados em arquivos de texto, onde cada amostra é armazenada em uma linha do arquivo, e as informações de cada amostra são separadas pelo caracter ";"(ponto e vírgula). Não foi utilizado um sistema gerenciador de banco de dados para armazenar as amostras, pois as mesmas não são alteradas durante a execução, são lidas uma única vez no início da execução.

Os dados históricos das variáveis selecionadas foram coletados de duas fontes: BM&FBovespa e EIA¹. Da BM&FBovespa foram coletados: o resumo diário de negociação das ações (preço de fechamento, preço de abertura, preço máximo, preço mínimo e volume comercializado), cotação de fechamento de índices de desempenho médio (IBovespa, IMAT, IFNC) e a taxa de câmbio do Dólar para Reais. Do EIA foi coletado a cotação do barril de petróleo *Brent*, informação esta que não estava disponível na BM&FBovespa.

Apesar de os dados, em sua maioria, terem sido coletados de uma mesma fonte, eles estavam armazenados em diferentes formatos e em diferentes tipos de arquivo. Para

¹ *Energy Information Administration* - órgão estatístico do governo dos Estados Unidos da América, responsável pela coleta, análise e disseminação de informações sobre energia.

facilitar a leitura destes dados pelo sistema previsor, foram gerados arquivos com os dados para cada uma das ações. Cada ação possui diversos arquivos e cada arquivo contém as informações necessárias para rodar um determinado experimento. Dessa forma, sempre que um experimento é configurado para executar, o sistema escolhe o arquivo correspondente ao experimento e à ação. Na [seção 5.3](#) estão descritos alguns dos experimentos executados.

Antes de serem armazenadas nos arquivos, as cotações do barril de petróleo *Brent* passaram por um pré-processamento já que sua precificação se dá em dólares. Para realizar a conversão, de dólares para reais, foi utilizada a taxa de câmbio coletada no *site* da BM&FBovespa.

Em uma análise preliminar dos dados coletados, verificou-se a ocorrência de cotações em datas diferentes entre os dados. Isso ocorre pelo fato de que a cotação do dólar e do preço do barril de petróleo *Brent* estão diretamente ligadas à mercados internacionais. Como a cotação ocorre apenas em dias úteis, há uma divergência nestas datas entre o mercado brasileiro e outros mercados. Para solucionar este problema de divergência de datas resolveu-se adotar as seguintes regras:

- Selecionar somente os dados com cotação em datas que correspondam à dias úteis no mercado brasileiro;
- Utilizar a última cotação existente (data imediatamente anterior) em caso de não haver cotação do dólar ou do barril de petróleo *Brent* em um dia útil do mercado brasileiro;
- Descartar cotações em dias não-úteis do mercado brasileiro.

Outro pré-processamento necessário foi a geração das informações dos indicadores de análise técnica escolhidos. Este pré-processamento pode ser caracterizado como uma extração de características, uma vez que são gerados novos dados a partir da série de fechamento das ações. As informações dos indicadores foram geradas da seguinte forma:

- Bandas de Bollinger: obtida através da configuração padrão, onde a média móvel simples é calculada sobre os últimos 20 dias. As bandas superior e inferior são formadas pelo acréscimo e decréscimo, respectivamente, de duas unidades do desvio-padrão dos últimos 20 dias. Portanto, o indicador Bandas de *Bollinger* gera duas informações;
- MACD: também modelado da forma mais tradicional, onde a linha MACD corresponde a diferença entre a MME de 12 dias e a MME de 26 dias, a linha de sinal é calculada através de uma MME de 9 dias sobre os dados da linha MACD e o histograma é a diferença entre a linha MACD e a linha de sinal. Portanto o indicador MACD é formado por 3 valores: linha MACD, linha de sinal e histograma;

- Intersecção de Médias Móveis: composta de duas médias móveis simples, uma de 13 dias considerada como média de curto prazo e outra de 20 dias sendo a média de longo prazo.

5.2 CONFIGURAÇÃO DAS REDES NEURAIS

Nesta seção são apresentados aspectos da configuração das redes neurais. Estes aspectos tratam da arquitetura de rede neural utilizada, parâmetros utilizados na rede, critério de parada do treinamento, dados de entrada e função de ativação.

Para implementar as redes neurais do trabalho proposto foi utilizada a linguagem de programação Java. Este protótipo de sistema predictor não possui uma interface gráfica para interação com o usuário. As configurações da execução podem ser alteradas através da edição de um arquivo XML.

5.2.1 Arquitetura

Uma RNA-MLP é composta por vários neurônios artificiais organizados em camadas. Porém, não há uma regra para a definição do número de neurônios de cada camada e do número de camadas escondidas. Esta escolha deve ocorrer de forma empírica (BORGES, 2012).

A camada de entrada de uma RNA-MLP tem seu tamanho definido de acordo com a natureza da aplicação (FREEMAN; SKAPURA, 1991). Em outras palavras, o número de neurônios da camada de entrada é definido de acordo com o número de dados de entrada que existem, cada dado é carregado para um neurônio da camada de entrada. Quanto maior é o vetor de entrada mais complexa é a rede neural que o utiliza. Por exemplo, um vetor 4-D (com 4 informações) gera uma rede neural menos complexa que um vetor de entrada 8-D.

Quanto às camadas escondidas, uma RNA-MLP pode ser construída com mais de uma. Freeman e Skapura (1991) afirmam que geralmente uma camada escondida é suficiente, mas que determinados problemas podem ser mais facilmente solucionados com mais de uma camada escondida. Para Hornik, Stinchcombe e White (1989) RNAs-MLP com uma camada escondida são capazes de gerar funções de aproximação universal muito precisas e satisfatórias. Neste trabalho as RNAs são construídas com 1 camada escondida. O número de neurônios desta camada é definido de forma empírica, pois não existe uma regra confiável para isso. Em cada experimento foram criadas e executadas diversas RNAs, com uma variação do número de neurônios nesta camada, visando escolher o melhor número de neurônios. Os testes iniciaram com 2 neurônios e, a cada execução, este número era incrementado de 2, testando até alcançar a marca de 40 neurônios nesta camada. Após estes testes obteve-se, de acordo com o desempenho de cada RNA, a melhor configuração

de neurônios na camada escondida para cada um dos experimentos.

Já a definição do número de neurônios da camada de saída, depende da quantidade de padrões de classificação. Neste trabalho existem duas possíveis classes: tendência de alta e tendência de baixa. Para aumentar a diferença entre os padrões de saída das duas classes, foi proposto a utilização de dois neurônios na camada de saída, definindo os padrões conforme exposto na [Figura 12](#). Também foi testada a utilização de apenas 1 neurônio na camada de saída, mas não obteve-se melhoras no desempenho do sistema predictor.

5.2.2 Inicialização dos Pesos e Parâmetros das RNAs

As conexões entre os neurônios de diferentes camadas são chamadas de sinapses, e em cada sinapse existe um peso sináptico associado. São estes pesos sinápticos que armazenam o conhecimento da RNA adquiridos durante o processo de treinamento. O treinamento consiste na atualização dos pesos sinápticos, de forma que, ao apresentar um vetor de entrada a rede gere o valor de saída desejado. De acordo com [Freeman e Skapura \(1991\)](#) estes pesos devem inicializados de forma aleatória, com valores pequenos, tipicamente no intervalo $[-0,5;0,5]$.

Uma prática comum na construção de RNAs-MLP é a utilização do parâmetro *bias* ([FREEMAN; SKAPURA, 1991](#)). O valor *bias* é aplicado externamente em todos os neurônios das camadas intermediárias e de saída, com o objetivo de alterar a entrada líquida da função de ativação. Ele pode ser considerado como uma sinapse extra, e seu peso é fixo. Normalmente, assim como foi adotado neste trabalho, o valor aplicado ao *bias* é igual a +1 ([HAYKIN, 2001](#)).

Outro parâmetro que deve ser definido em RNAs-MLP, a taxa de aprendizagem, tem um efeito significativo na performance da RNA. O valor da taxa de aprendizagem normalmente deve ser pequeno, no intervalo $[0,05;0,25]$ ([FREEMAN; SKAPURA, 1991](#)). Quanto menor for o parâmetro da taxa de aprendizagem menores serão as atualizações dos pesos sinápticos, de uma iteração para outra, podendo levar a um treinamento muito lento, sendo necessárias várias iterações para finalizar o treinamento. Já um valor muito grande nesta taxa, apesar de acelerar o processo de treinamento, pode tornar a rede instável ([HAYKIN, 2001](#)). Neste trabalho a taxa de aprendizagem foi definida como 0,15.

5.2.3 Critério de Parada do Treinamento

Segundo [Haykin \(2001\)](#), não se pode demonstrar que o algoritmo *backpropagation* converge e não existem critérios de parada bem definidos para encerrar o processo de treinamento. No entanto, existem alguns critérios razoáveis para encerrar o ajuste dos pesos. Neste trabalho foram implementados três critérios de parada: limite de iterações, variação mínima do erro quadrático médio e validação cruzada.

O critério de parada limite de iterações consiste em estabelecer um limite máximo de iterações do treinamento. Este critério sempre é utilizado aliado a outro, e serve como garantia que o treinamento irá encerrar, uma vez que os outros critérios podem não ser atingidos durante o treinamento. Inicialmente definiu-se o número de 10.000 iterações como limite, mas ao analisar os relatórios de execução verificou-se que a maioria das execuções alcançava o critério de parada antes de 1.000 iterações. Dessa forma fixou-se em 1.000 o limite de iterações no treinamento.

O critério de parada baseado na variação mínima do erro quadrático médio considera que o algoritmo *backpropagation* converge quando a taxa absoluta de variação do erro médio quadrado por iteração for suficientemente pequena (HAYKIN, 2001). Para calcular esta taxa é necessário calcular o valor instantâneo do erro $\xi(n)$, através de

$$\xi(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (5.1)$$

onde j é um neurônio pertencente a camada de saída C e $e_j^2(n)$ corresponde ao sinal de erro do neurônio j na iteração n conforme definido na [Equação 3.5](#). O erro quadrático médio em uma iteração n é definido como

$$\xi_{med} = \frac{1}{N} \sum_{n=1}^N \xi(n). \quad (5.2)$$

O treinamento é encerrado quando a variação do erro quadrático médio entre duas iterações for menor que 0,0001.

Outro critério de parada implementado foi a utilização da validação cruzada através do método de treinamento com parada antecipada. Em situações típicas o erro quadrático médio decresce com o aumento do número de iterações (HAYKIN, 2001). Tendo como objetivo uma boa generalização é muito difícil perceber o momento certo para encerrar o treinamento. Caso o treinamento se estenda demais, a generalização pode ficar comprometida, uma vez que o sistema pode "decorar" o comportamento das amostras de treinamento. Através da validação cruzada pode-se identificar o início do excesso de treinamento. O conjunto de treinamento é dividido em conjunto de estimação e conjunto de validação. Após determinados períodos de estimação (treinamento) são calculados os erros de validação através do cálculo do erro quadrático médio. Quando o erro de validação começar a crescer é o momento certo para encerrar o treinamento, ou seja, a partir deste instante o treinamento estará perdendo a capacidade de generalização e começando a aprender demais o conjunto das amostras de treinamento. Para separar o conjunto de validação que é equivalente a um percentual do número de amostras do conjunto de treinamento, calculado através de

$$R = 1 - \frac{\sqrt{2 * W - 1} - 1}{\sqrt{2 * (W - 1)}} \quad (5.3)$$

onde W representa a soma de todos os parâmetros livres da RNA (pesos sinápticos e *bias*), R representa o percentual de amostras que serão alocadas para o conjunto de validação, o restante $(1 - R)$ corresponde ao percentual de amostras do conjunto de estimação.

Foram realizados vários experimentos para analisar os resultados obtidos com cada um dos critérios de parada implementados. Notou-se que o critério de parada com validação cruzada encerrava os treinamentos de forma muito precoce, pois os resultados nos testes eram inferiores se comparados aos resultados obtidos com o outro critério. Analisando a execução do treinamento, passo-a-passo, notou-se que o erro quadrático médio oscila entre as iterações, sendo esta uma situação atípica, onde o erro não apenas decresce durante o treinamento. Por este motivo, os critérios de parada utilizados nos treinamentos posteriores foram a variação mínima do erro quadrático médio e o limite máximo de iterações.

5.2.4 Função de Ativação

A função de ativação tem como objetivo restringir os valores de saída de cada neurônio em um intervalo fixo, impedindo que um neurônio envie para outro um valor muito grande ou pequeno. Caso não fosse utilizada, o erro produzido em um neurônio de saída poderia ser grande, dificultando o ajuste dos pesos sinápticos, e conseqüentemente o processo de treinamento.

A escolha da função de ativação depende de como os dados de saída são representados (FREEMAN; SKAPURA, 1991). Neste trabalho as possíveis saídas da RNA são $[-1; 1]$ para tendência de baixa e $[1; -1]$ para tendência de alta. Considerando esta abrangência de valores, escolheu-se a função de ativação tangente hiperbólica, que atende aos requisitos (Figura 10).

5.2.5 Normalização dos Dados de Entrada

A normalização dos dados de entrada é um pré-processamento fundamental em aplicações de redes neurais artificiais (KHAN; BANDOPADHYAYA; SHARMA, 2008). É necessária pois ao apresentar valores de grande escala para a função de ativação ela restringirá a saída dos valores. No caso da função de tangente hiperbólica, qualquer valor acima de 1 apresentado na entrada da função terá como saída o valor 1. Muita informação é perdida quando os dados de entrada não são normalizados.

Os dados de entrada são normalizados conforme os limites inferior e superior da função de ativação (MACIEL; BALLINI, 2009). Assim os valores foram normalizados para o intervalo $[-1; 1]$, que são os limites inferior e superior de abrangência da função tangente hiperbólica. Esta normalização foi obtida através de

$$x_n = 2 * (x - \min_{\mathbf{x}}) / (\max_{\mathbf{x}} - \min_{\mathbf{x}}) - 1 \quad (5.4)$$

onde x_n representa o valor normalizado para o intervalo $[-1; 1]$, x é o valor a ser normalizado, \min_x é o menor valor do vetor x e \max_x é o maior valor presente no vetor x , limitando o vetor ao conjunto de amostras de treinamento. A normalização foi aplicada sobre todas as séries temporais de entrada.

5.3 EXPERIMENTOS

Para a realização dos experimentos, os dados coletados foram particionados em dois conjuntos: treinamento e teste. O conjunto de treinamento corresponde a $\frac{2}{3}$ dos dados, enquanto o conjunto de teste é composto pelo $\frac{1}{3}$ restante. É importante lembrar que apenas através do conjunto de treinamento ocorre a correção dos pesos sinápticos. O conjunto de teste será utilizado somente para avaliar o desempenho do sistema.

A maioria dos algoritmos de aprendizagem, incluindo as RNAs-MLP, espera uma distribuição equilibrada do número de amostras de cada classe (HE; GARCIA, 2009). A apresentação de um conjunto desbalanceado de amostras pode prejudicar a concepção de um classificador. Para evitar este problema, os conjuntos de treinamento e teste foram balanceados, eliminando aleatoriamente amostras da classe que tinha um maior número de amostras. Esta regra do balanceamento é desrespeitada quando se está calculando a simulação de investimento, pois para isso é necessária a avaliação de todas as amostras do período testado.

Uma boa prática nas tarefas de classificação de padrões, segundo Haykin (2001), é tornar aleatória a ordem de apresentação das amostras em cada iteração do treinamento. Dessa forma evita-se que amostras sucessivas apresentadas a RNA pertençam a uma mesma classe, o que poderia dificultar o treinamento do classificador, afetando sua capacidade de generalização.

Como as RNAs são inicializadas com pesos aleatórios em suas conexões sinápticas, pode ocorrer que a RNA não consiga convergir no seu treinamento. Por este motivo, cada experimento foi realizado várias vezes, e somente o melhor destes foi selecionado. Para seleção do melhor experimento de cada modelo foram definidas duas regras:

- Ter uma distribuição minimamente equilibrada na previsão de tendência, de modo que

$$DPT = VP/(VP + VN) \quad (5.5)$$

onde DPT corresponde ao percentual de Distribuição na Previsão de Tendência, VP é o verdadeiro positivo ou número de acertos na previsão de tendência de alta, VN é o verdadeiro negativo ou número de acertos na previsão de tendência de baixa. Para ter uma distribuição minimamente equilibrada o experimento deve apresentar o valor DPT entre 0.3 e 0.7. Essa restrição evita que sejam escolhidos experimentos

que acertam apenas tendência de alta ou de baixa, que prevêm sempre a mesma tendência independente dos dados de entrada.

- Possuir a maior taxa de acerto dentre os experimentos que satisfazem a regra anterior.

As subseções a seguir mostram os resultados obtidos nos diversos testes realizados no sistema entre as bases de dados. O teste binomial para diferenças de proporção (MOTTA, 2006) é usado para testar se as diferenças de duas proporções binomiais são estatisticamente significantes. Quando não especificado, o nível de significância α é igual a 0,05. Este teste de significância é aplicado apenas sobre as taxas de acerto.

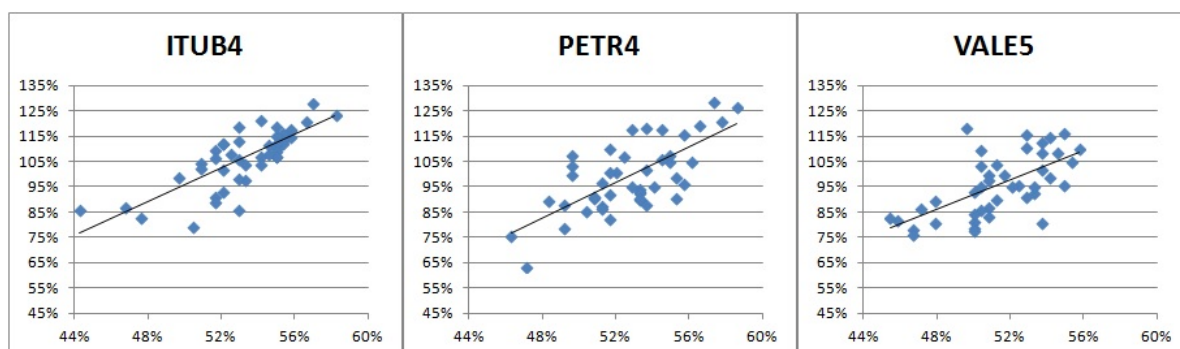
Conforme a arquitetura de sistema predictor proposta (Figura 12), os experimentos foram divididos em RNA com Variáveis Dependentes e RNA com Variáveis Exógenas e Indicadores. As subseções que seguem relatam os experimentos realizados para ambas as RNAs.

5.3.1 RNA com Variáveis Dependentes

A RNA com variáveis dependentes é caracterizada pela utilização dos valores que resumem o dia de negociação de uma ação em bolsa como dados de entrada. Os experimentos foram realizados com a utilização das variáveis dependentes em sua forma estática e em uma forma dinâmica, obtida através do cálculo da diferença relativa dos últimos dias.

Os experimentos são avaliados por duas métricas: taxa de acerto e percentual de retorno de uma simulação de investimento. No entanto, as taxas obtidas por essas métricas não são perfeitamente correlacionadas. O coeficiente de correlação ρ foi calculado a partir de todo o conjunto de experimentos realizados. Para a ação ITUB4 $\rho = 0.77$, para PETR4 $\rho = 0.70$ e para VALE5 $\rho = 0.68$. Isto significa que a taxa de acerto apresenta uma forte correlação com o retorno da simulação de investimento, mas como essa correlação não é perfeita ($\rho = 1$) podem ocorrer situações em que a taxa de acerto aumenta e o retorno com a simulação de investimento diminui. O gráfico de dispersão dos experimentos realizados mostra essa correlação entre as variáveis (Figura 13).

Figura 13: Gráfico de dispersão da relação entre taxa de acerto e taxa de retorno da simulação de investimento, obtido através dos experimentos realizados na RNA com Variáveis Dependentes.



Fonte: desenvolvido pelo autor

5.3.1.1 Variáveis de Entrada Estáticas

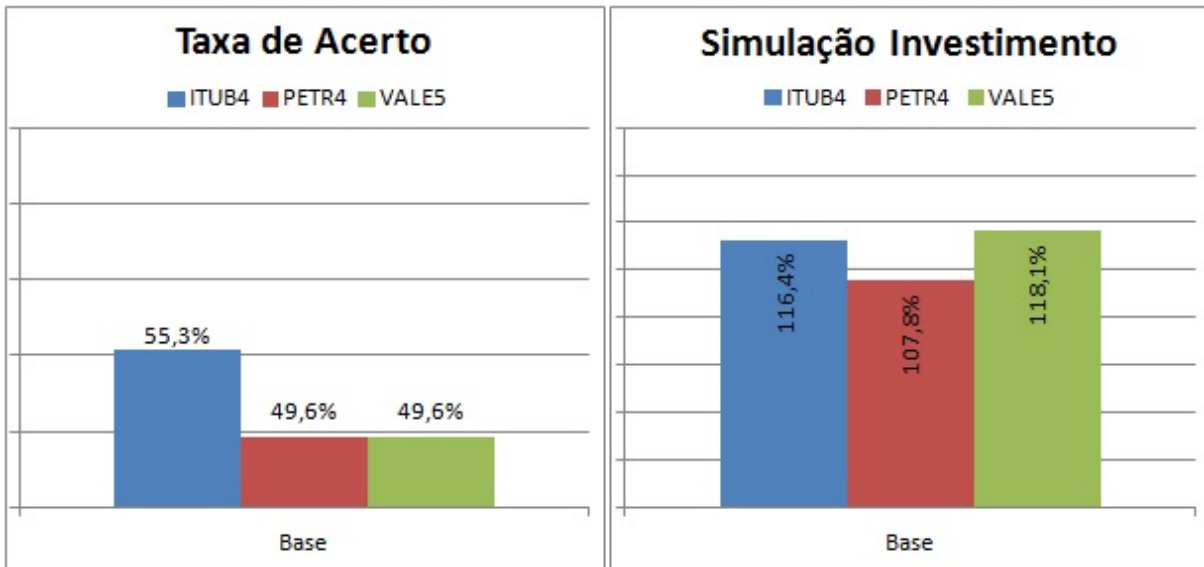
Os experimentos foram realizados utilizando variáveis dependentes, em sua forma estática, como valores de entrada das RNAs. Os dados estáticos correspondem aos valores reais de cotação, não ocorre transformação prévia dos dados. Através das variáveis descritas na Tabela 2 foram gerados diversos modelos para experimentos, com diferentes combinações de variáveis como dados de entrada das RNAs.

Tabela 2: Descrição das variáveis dependentes do ativo, em sua forma estática.

Descrição	Variável de Entrada
Fechamento da ação no período $t - x$	$vFechAcao(t - x)$
Abertura da ação no período $t - x$	$vAbeAcao(t - x)$
Valor máximo da ação no período $t - x$	$vMaxAcao(t - x)$
Valor mínimo da ação no período $t - x$	$vMinAcao(t - x)$
Volume negociado da ação no período $t - x$	$vVolAcao(t - x)$

Para iniciar os experimentos foi definido um modelo Base, que utiliza o vetor de características $[vFechAcao(t), vAbeAcao(t), vMaxAcao(t), vMinAcao(t), vVolAcao(t)]$ como entrada da RNA, onde o período de tempo t refere-se ao dia atual. Nota-se que estas variáveis representam todo o conjunto de variáveis dependentes, e o vetor de entrada tem uma dimensionalidade 5-D. Os resultados deste experimento apresentaram um percentual de acerto não melhor que a chance, como mostra a Figura 14. No que se refere à simulação de investimento com o modelo Base, obteve-se lucro com as 3 ações analisadas, sendo 16,4% na ação ITUB4, 7,8% na ação PETR4 e 18,1% na VALE5. Isso representa um resultado satisfatório, pois a inflação acumulada para o período analisado foi de aproximadamente 6%.

Figura 14: Taxa de acerto e simulação de investimento do modelo Base da RNA com variáveis dependentes, utilizando dados estáticos.



Fonte: desenvolvido pelo autor

Como o experimento com o modelo Base não produziu taxas de acerto melhores que a chance, dois novos modelos foram construídos isolando as variáveis do modelo Base: a série temporal de fechamento da ação ($vFechAcao(t)$) e as séries temporais com as outras informações dependentes ($vAbeAcao(t)$, $vMaxAcao(t)$, $vMinAcao(t)$, $vVolAcao(t)$). Estes modelos foram denominados como Fechamento, cujo vetor de entrada possui uma dimensionalidade 1-D, e Outros, cujo vetor tem uma dimensão 4-D. As RNAs construídas utilizando estes vetores como entrada não obtiveram resultados melhores. Portanto, as informações utilizadas no modelo Base se complementam e são capazes de produzir melhores resultados se utilizadas em conjunto do que se utilizadas isoladamente.

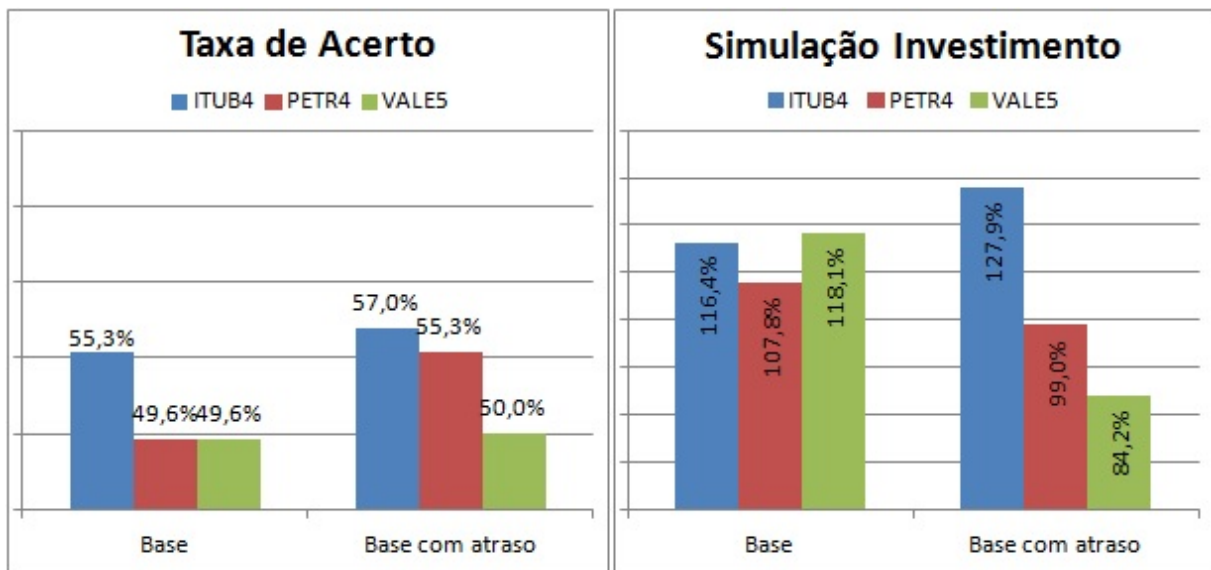
Como os modelos Fechamento e Outros não melhoraram o desempenho, e com o intuito de melhorar os resultados obtidos, criou-se um novo modelo denominado Base com Atraso, o qual possui um vetor de entrada de dimensão 15-D, composto por todas variáveis dependentes, com a cotação do dia atual e dos dois dias úteis imediatamente anteriores. O vetor de entrada utilizado neste experimento pode ser expresso como

$$\text{Base com atraso} = [vFechAcao(t-x), vAbeAcao(t-x), vMaxAcao(t-x), vMinAcao(t-x), vVolAcao(t-x)] \quad (5.6)$$

onde $x \in [0; 2]$. Apesar deste experimento produzir taxas de acerto superiores ao modelo Base, como mostra a [Figura 15](#), esta diferença não é estatisticamente significativa. Entretanto, esta diferença proporcionou um melhor retorno na simulação de investimento da ação ITUB4. O mesmo não ocorreu com as outras ações, que tiveram uma queda

no percentual de retorno e apresentaram prejuízo no período analisado. Essa falta de correlação entre taxa de acerto e percentual de retorno financeiro ocorre pois, em algumas situações, acertar a tendência é mais vantajoso que em outras. Por exemplo, considerando que o investidor tenha ações para vender e que o sistema acerte a tendência de baixa antes de uma queda de 10%. Esse acerto de tendência é mais vantajoso que acertar a tendência de baixa antes de uma queda de 1%. Portanto, melhorar a taxa de acerto não dá garantias de melhores retornos financeiros.

Figura 15: Taxa de acerto e simulação de investimento dos modelos Base e Base com atraso da RNA com variáveis dependentes, utilizando dados estáticos.



Fonte: desenvolvido pelo autor

Além dos citados, foram realizados diversos experimentos com dados estáticos, gerando diferentes combinações das variáveis dependentes do ativo, para servirem de entrada na RNA. Todos esses experimentos não produziram melhores resultados.

5.3.1.2 Variáveis de Entrada Dinâmicas

Como os experimentos realizados com as variáveis de entrada estáticas não produziram resultados melhores que a chance, resolveu-se explorar a capacidade de generalização das RNAs através da utilização de dados dinâmicos. Os dados dinâmicos são caracterizados por representarem um comportamento de variação dos dados, como por exemplo a diferença de valor dos y últimos dias, ou a taxa de variação entre o dia x e o dia y . Neste trabalho, os dados dinâmicos foram gerados através do cálculo de diferença relativa entre os últimos y dias.

As variáveis dependentes do ativo, na sua forma dinâmica, estão descritas na [Tabela 3](#). Estas variáveis foram utilizadas para a criação dos modelos de dados para os experimentos da RNA com variáveis dependentes.

Tabela 3: Descrição das variáveis dependentes do ativo, em sua forma dinâmica, geradas através do cálculo da diferença relativa.

Descrição	Variável de Entrada
Diferença do $vFechAcao(t - x)$ entre os últimos y dias	$dFechAcao(t - x, y)$
Diferença do $vAbeAcao(t - x)$ entre os últimos y dias	$dAbeAcao(t - x, y)$
Diferença do $vMaxAcao(t - x)$ entre os últimos y dias	$dMaxAcao(t - x, y)$
Diferença do $vMinAcao(t - x)$ entre os últimos y dias	$dMinAcao(t - x, y)$
Diferença do $vVolAcao(t - x)$ entre os últimos y dias	$dVolAcao(t - x, y)$

Em um primeiro momento, os dados dinâmicos foram gerados através do cálculo da diferença relativa dos últimos 2 dias. A fórmula utilizada para calcular a diferença relativa entre os últimos 2 dias foi definida como

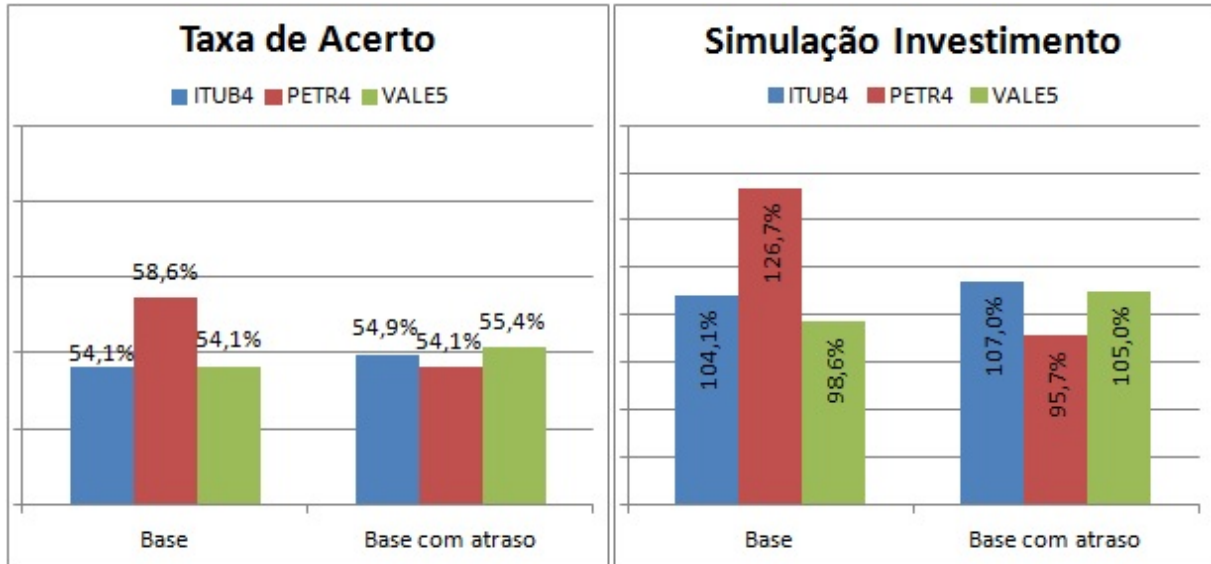
$$dVar(t, 2) = vVar(t) - vVar(t - 1) \quad (5.7)$$

onde t representa o dia atual e $vVar$ representa o valor estático da variável.

Os experimentos realizados com dados dinâmicos seguiram a mesma lógica dos experimentos realizados com dados estáticos. Inicialmente foi definido um modelo Base da RNA com variáveis dinâmicas dependentes, utilizando o vetor de características $[dFechAcao(t, 2), dAbeAcao(t, 2), dMaxAcao(t, 2), dMinAcao(t, 2), dVolAcao(t, 2)]$ como entrada para a rede neural, onde t é o dia atual. Este modelo produziu resultados semelhantes aos apresentados pelo modelo Base com dados estáticos, com uma taxa de acerto um pouco mais elevada, como mostra a [Figura 16](#). Destaca-se o desempenho obtido com este modelo na ação PETR4, que alcançou uma taxa de acerto de 58,6% e um lucro de 26,7% na simulação de investimento.

Para verificar o comportamento das RNAs com as variáveis isoladas, foram gerados os modelos Fechamento ($dFechAcao(t, 2)$) e Outros ($dAbeAcao(t, 2)$, $dMaxAcao(t, 2)$, $dMinAcao(t, 2)$ e $dVolAcao(t, 2)$). Estes experimentos não produziram resultados melhores que o modelo Base. Além destes, foi criado o modelo Base com Atraso, com as mesmas variáveis do modelo Base porém com dados do dia atual e dos últimos 2 dias, cujo vetor de características tem uma dimensionalidade 15-D. Este experimento se mostrou consistente, pois todas as ações produziram resultados semelhantes, tanto para a taxa de acerto quanto para a simulação financeira calculada, porém nenhum deles apresentou um retorno financeiro muito maior que a inflação acumulada no período. O retorno financeiro para as ações ITUB4 e VALE5 no experimento Base com atraso foi maior se comparado com o modelo Base, como mostra a [Figura 16](#). Em contrapartida, para a ação PETR4 houve queda de mais de 30% no retorno financeiro calculado entre o modelo Base e o modelo Base com atraso.

Figura 16: Taxa de acerto e simulação de investimento dos modelos Base e Base com atraso da RNA com variáveis dependentes, utilizando dados dinâmicos calculados pela diferença relativa de 2 dias.



Fonte: desenvolvido pelo autor

Foram testadas várias outras combinações de variáveis dinâmicas dependentes, com uma diferença relativa de 2 dias, como entrada na RNA. Nenhuma dessas outras combinações apresentou resultados melhores dos que já foram apresentados.

Como os dados dinâmicos gerados pela diferença relativa de 2 dias alcançaram resultados semelhantes, e em alguns casos melhores que os estáticos, resolveu-se testar o poder de generalização das RNAs com dados dinâmicos gerados através do cálculo da diferença relativa dos últimos 3 e 5 dias. A diferença relativa para mais de 2 dias é calculada utilizando regressão polinomial ortogonal de primeira ordem na trajetória temporal da característica (SOONG; ROSENBERG, 1988). O coeficiente polinomial ortogonal de primeira ordem, ou a inclinação da reta generalizada, para 3 e 5 dias possui a seguinte forma:

$$dVar(t, 3) = vVar(t) * 0.5 + vVar(t - 1) * 0 + vVar(t - 2) * (-0.5) \quad (5.8)$$

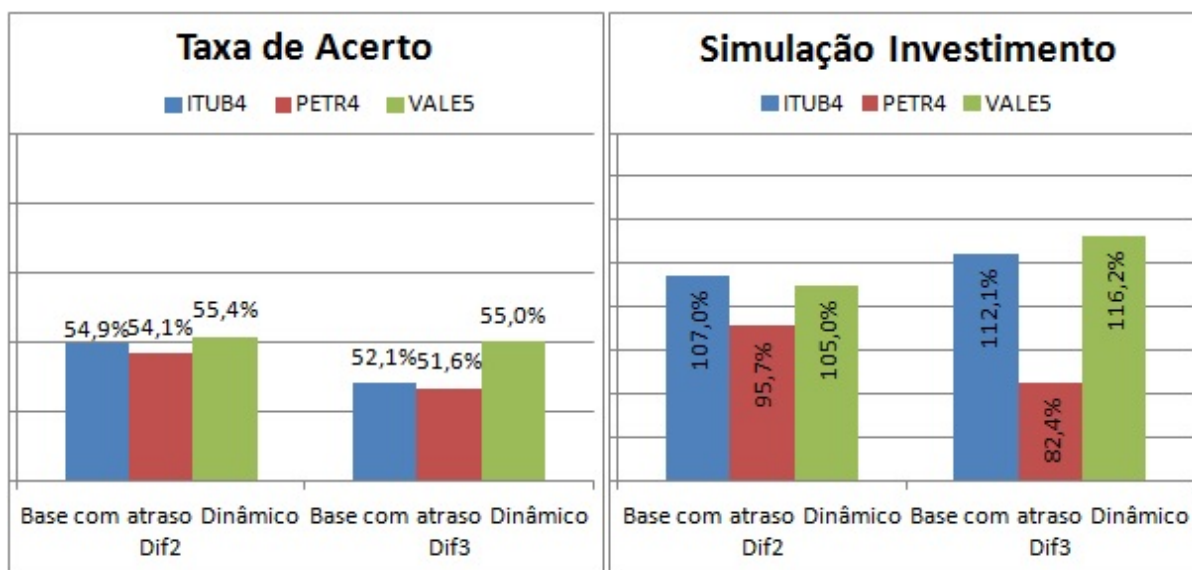
e

$$dVar(t, 5) = vVar(t) * 0.2 + vVar(t - 1) * 0.1 + vVar(t - 2) * 0 + vVar(t - 3) * (-0.1) + vVar(t - 4) * (-0.2) \quad (5.9)$$

onde t representa o dia atual, $vVar$ representa o valor estático da variável, $dVar(t, 3)$ corresponde a diferença relativa de 3 dias e $dVar(t, 5)$ representa a diferença relativa de 5 dias.

Dos experimentos realizados com dados dinâmicos calculados pela diferença relativa dos últimos 3 dias destaca-se o modelo Base com atraso, que contém as mesmas variáveis do modelo Base com atraso anterior, porém a diferença calculada é de 3 dias. Este modelo, como mostra a [Figura 17](#), apesar de apresentar taxas de acerto menores nas 3 ações analisadas, possibilitou um ganho financeiro em 2 dessas ações conforme simulação de investimento realizada.

Figura 17: Taxa de acerto e simulação de investimento dos modelos Base com atraso da RNA com variáveis dependentes, utilizando dados dinâmicos calculados pela diferença dos últimos 2 e 3 dias.



Fonte: desenvolvido pelo autor

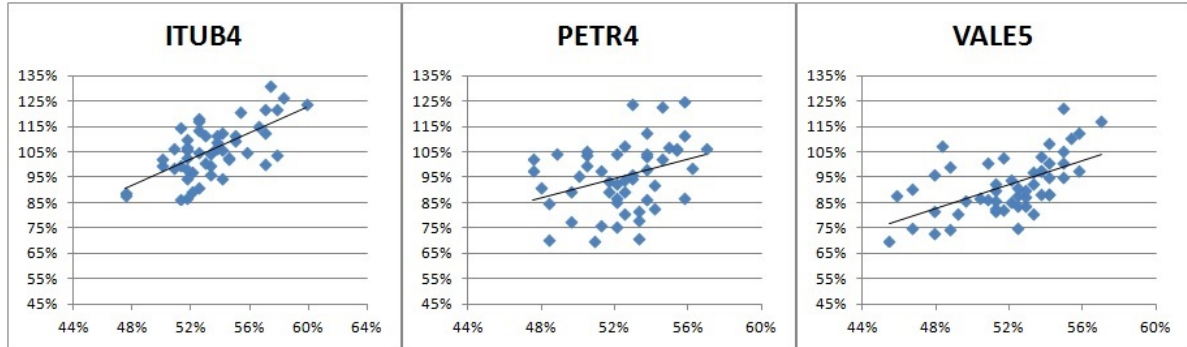
Por sua vez, os experimentos realizados com a diferença relativa de 5 dias sempre apresentavam taxas inferiores em relação aos realizados com dados estáticos, dinâmicos da diferença de 2 dias e dinâmicos da diferença de 3 dias. Com isso pode-se concluir que a informação da diferença entre 2, ou no máximo 3 dias, é suficiente para o modelo classificador e que agregar a informação da diferença de mais dias não auxilia na classificação.

5.3.2 RNA com Variáveis Exógenas e Indicadores

Os experimentos realizados na RNA com variáveis exógenas e indicadores, assim como nos realizados com a RNA com variáveis dependentes, são avaliados pelas métricas taxa de acerto e percentual de retorno obtido através de uma simulação de investimento. A partir do conjunto de experimentos realizados na RNA com variáveis exógenas e indicadores, foi possível verificar que existe uma correlação moderada entre a taxa de acerto e o retorno financeiro. Os coeficientes de correlação para esta RNA foram menores se comparados com os obtidos com os experimentos da RNA com variáveis dependentes. Para a ação ITUB4

$\rho = 0.62$, para a PETR4 $\rho = 0.39$ e para a VALE5 $\rho = 0.49$. Esta correlação moderada pode ser verificada no gráfico de dispersão apresentado na Figura 18.

Figura 18: Gráfico de dispersão da relação entre taxa de acerto e taxa de retorno do investimento simulado, obtido através dos experimentos realizados na RNA com Variáveis Exógenas e Indicadores.



Fonte: desenvolvido pelo autor

5.3.2.1 Variáveis de Entrada Estáticas

Inicialmente os experimentos da RNA com variáveis exógenas e indicadores utilizaram os dados em sua forma estática. A descrição das variáveis utilizadas para a realização destes experimentos é exibida da Tabela 4. Através da combinação deste conjunto de variáveis foram gerados os modelos de dados de entrada da RNA para os experimentos.

Tabela 4: Descrição das variáveis exógenas e de indicadores, em sua forma estática.

Descrição	Variável de Entrada
Fechamento do IBovespa no período $t - x$	$vIbov(t - x)$
Fechamento do índice setorial da ação no período $t - x$	$vIndSet(t - x)$
Fechamento da cotação do dólar no período $t - x$	$vDolar(t - x)$
Valor da linha MACD no período $t - x$	$vMacd(t - x)$
Valor da linha de sinal MACD no período $t - x$	$vMacdSinal(t - x)$
Valor do histograma MACD no período $t - x$	$vMacdHist(t - x)$
Valor inferior das Bandas de Bollinger no período $t - x$	$vBBInf(t - x)$
Valor superior das Bandas de Bollinger no período $t - x$	$vBBSup(t - x)$
Valor da MMS de longo prazo no período $t - x$	$vMMSMaior(t - x)$
Valor da MMS de curto prazo no período $t - x$	$vMMSMenor(t - x)$

Seguindo a mesma lógica dos experimentos realizados na RNA com variáveis dependentes, definiu-se um modelo Base para os experimentos realizados na RNA com variáveis exógenas e indicadores. O modelo Base utiliza o vetor de características $[vIbov(t), vIndSet(t), vDolar(t), vMacd(t), vMacdSinal(t), vMacdHist(t), vBBInf(t), vBBSup(t), vMMSMaior(t), vMMSMenor(t)]$ como entrada na RNA, onde t corresponde ao dia atual.

Em seguida, foram gerados quatro modelos isolando as variáveis exógenas e indicadores. Estes modelos utilizam os seguintes vetores de características como entrada na RNA:

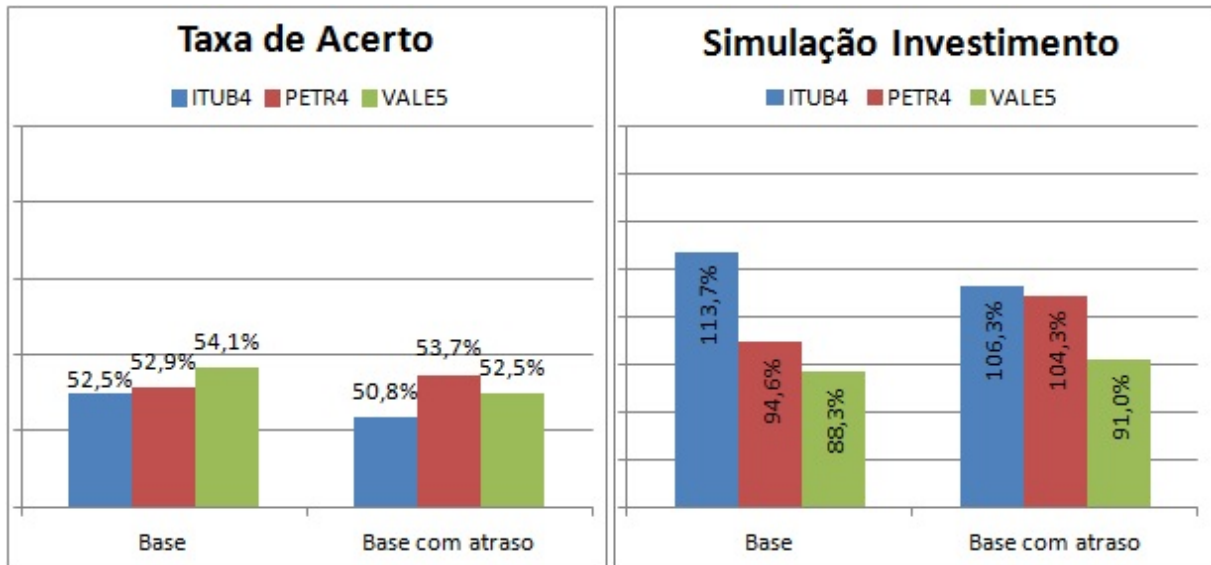
- Exógenos = $[vIbov(t), vIndSet(t), vDolar(t)]$
- Macd = $[vMacd(t), vMacdSinal(t), vMacdHist(t)]$
- Bollinger = $[vBBInf(t), vBBSup(t)]$
- Cruzamento MMS = $[vMMSMaior(t), vMMSMenor(t)]$

As RNAs construídas com os vetores de características dos modelos Exógenos, Macd, Bollinger e Cruzamento MMS não apresentaram taxas de acerto melhores que a chance. No que se refere aos percentuais de retorno obtidos através da simulação com estes modelos, nenhum deles possibilitou a obtenção de lucro. O modelo Base, apesar de não apresentar taxa de acerto melhor que a chance, possibilitou a obtenção de uma taxa de retorno na simulação de investimento de 113,7% na ação ITUB4. Para as ações PETR4 e VALE5 a taxa de retorno foi menor que 100%, representando um prejuízo no investimento.

Após estes experimentos, testou-se o modelo Base com Atraso para a RNA com variáveis exógenas e indicadores. Este modelo utilizou um vetor de características com dimensionalidade 30-D, o qual possui o conjunto de todas as variáveis exógenas e indicadores, com os valores de cotação do dia atual e dos dois dias úteis imediatamente anteriores. Este experimento produziu valores de retorno no investimento maiores que 100% em 2 das 3 ações, como mostra a [Figura 19](#), obtendo um retorno maior em todas as ações se comparado ao modelo Base.

Vários outros experimentos foram testados, combinando as variáveis exógenas e indicadores na entrada da RNA, porém nenhum apresentou um desempenho melhor. Estes outros experimentos realizados incluíram a combinação de dados de diferentes indicadores, de indicadores com variáveis exógenas e variação nos dados exógenos utilizados.

Figura 19: Taxa de acerto e simulação de investimento dos modelos Base e Base com Atraso da RNA com variáveis exógenas e indicadores, utilizando dados estáticos.



Fonte: desenvolvido pelo autor

5.3.2.2 Variáveis de Entrada Dinâmicas

Com o objetivo de verificar o poder de generalização da RNA com variáveis exógenas e indicadores e de alcançar melhores taxas de acerto e de retorno no investimento, foram realizados experimentos com dados dinâmicos. A descrição das variáveis utilizadas para a criação dos modelos com dados dinâmicos é exibida na [Tabela 5](#).

Tabela 5: Descrição das variáveis exógenas e indicadores, em sua forma dinâmica, geradas através do cálculo da diferença relativa.

Descrição	Variável de Entrada
Diferença do $vIbov(t-x)$ entre os últimos y dias	$dIbov(t-x, y)$
Diferença do $vIndSet(t-x)$ entre os últimos y dias	$dIndSet(t-x, y)$
Diferença do $vDolar(t-x)$ entre os últimos y dias	$dDolar(t-x, y)$
Diferença do $vMacd(t-x)$ entre os últimos y dias	$dMacd(t-x, y)$
Diferença do $vMacdSinal(t-x)$ entre os últimos y dias	$dMacdSinal(t-x, y)$
Diferença do $vMacdHist(t-x)$ entre os últimos y dias	$dMacdHist(t-x, y)$
Diferença do $vBBInf(t-x)$ entre os últimos y dias	$dBBInf(t-x, y)$
Diferença do $vBBSup(t-x)$ entre os últimos y dias	$dBBSup(t-x, y)$
Diferença do $vMMSMaior(t-x)$ entre os últimos y dias	$dMMSMaior(t-x, y)$
Diferença do $vMMSMenor(t-x)$ entre os últimos y dias	$dMMSMenor(t-x, y)$

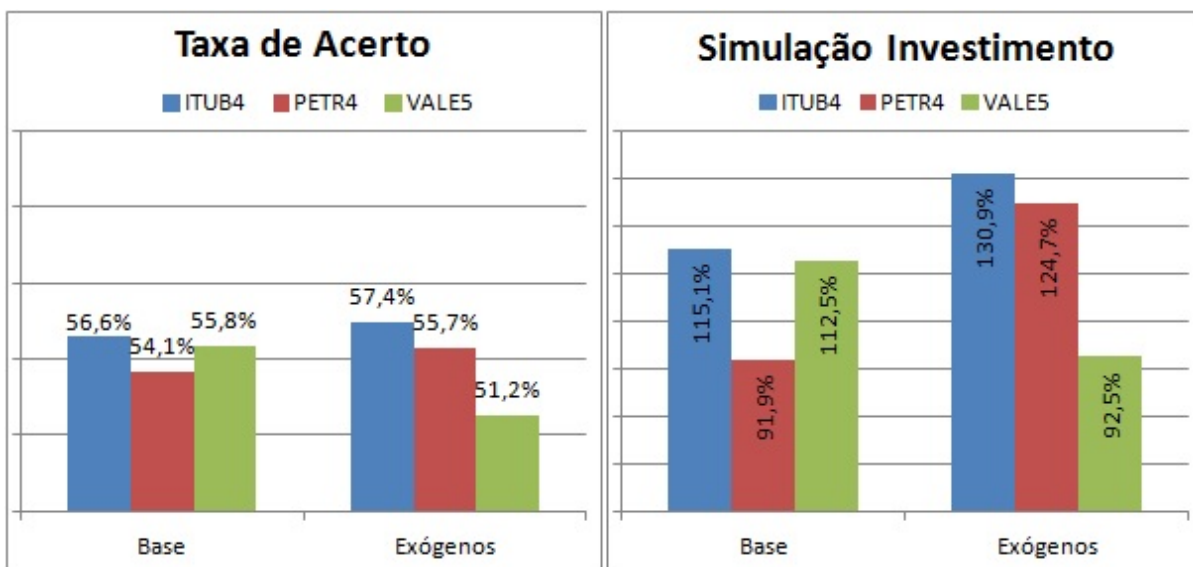
Os dados dinâmicos para estes experimentos foram obtidos através do cálculo da diferença relativa de 2 dias, conforme a [Equação 5.7](#). Os vetores de características dos modelos testados são:

- Base = $[dIbov(t, 2), dIndSet(t, 2), dDolar(t, 2), dMacd(t, 2), dMacdSinal(t, 2), dMacdHist(t, 2), dBBIInf(t, 2), dBBSup(t, 2), dMMSMaior(t, 2), dMMSMenor(t, 2)]$
- Exógenos = $[dIbov(t, 2), dIndSet(t, 2), dDolar(t, 2)]$
- Macd = $[dMacd(t, 2), dMacdSinal(t, 2), dMacdHist(t, 2)]$
- Bollinger = $[dBBIInf(t, 2), dBBSup(t, 2)]$
- Cruzamento MMS = $[dMMSMaior(t, 2), dMMSMenor(t, 2)]$

Os resultados obtidos com estes modelos não foram melhores que os resultados alcançados com dados estáticos. Portanto, para a RNA com variáveis dependentes e indicadores, dados dinâmicos calculados pela diferença relativa dos últimos 2 dias apresentam um desempenho inferior se comparados com dados estáticos.

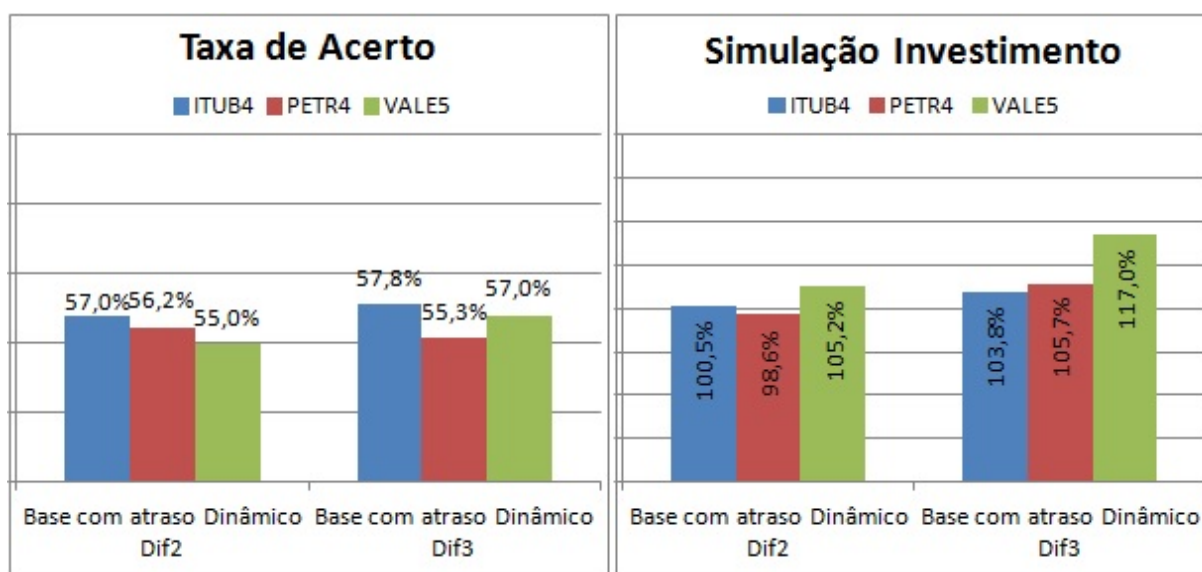
Mesmo com o indicativo que dados dinâmicos resultam em RNAs com desempenho inferior que dados estáticos quando a RNA é formada por dados exógenos e de indicadores, resolveu-se testar estes mesmos modelos, porém com dados dinâmicos calculados através da diferença de 3 dias. Os modelos Base e Exógenos se destacaram alcançando boas taxas de retorno de investimento, como mostra a [Figura 20](#). Destaca-se a taxa de retorno de investimento obtida através do modelo Exógenos sobre a ação ITUB4 que obteve um lucro de 30,9% no período simulado.

Figura 20: Taxa de acerto e simulação de investimento dos modelos Base e Exógenos da RNA com variáveis dependentes e indicadores, utilizando dados dinâmicos calculados pela diferença relativa de 3 dias.



Foram testados ainda os modelos Base com atraso com dados dinâmicos, calculados pela diferença relativa de 2 e de 3 dias. Ambos os modelos alcançaram taxas de acerto ligeiramente superiores se comparadas as obtidas nos modelos anteriores. Os percentuais de retorno de uma simulação de investimento também alcançaram valores que implicariam em um lucro no período analisado, conforme é exibido na [Figura 21](#).

Figura 21: Taxa de acerto e simulação de investimento dos modelos Base com da RNA com variáveis exógenas e indicadores, utilizando dados dinâmicos calculados pela diferença de 2 e de 3 dias.



Fonte: desenvolvido pelo autor

Além dos experimentos com os modelos citados, foram gerados diversos outros modelos com dados dinâmicos calculados pela diferença de 2 e de 3 dias. Também foram testados os modelos com dados formados pela diferença relativa dos últimos 5 dias. Nenhum destes outros experimentos apresentou melhores resultados que os já citados.

5.4 FUSÃO

Visando melhorar os resultados obtidos foram realizados alguns testes de fusão. Estes testes podem ser divididos em dois tipos quanto a natureza da fusão: fusão das características e fusão dos sistemas especialistas.

5.4.1 Fusão das Características

A fusão das características corresponde a união das variáveis de entrada, visando testar a complementariedade das informações, antes testadas separadamente, para gerar resultados melhores. Foram testadas 2 formas de fusão de características: fusão de dados

estáticos com dados dinâmicos e fusão das variáveis dependentes com as variáveis exógenas e indicadores.

A fusão dos dados estáticos com os dados dinâmicos não produziu resultados melhores que os obtidos somente com dados estáticos ou dinâmicos. Esta fusão somente trouxe uma complexidade maior para a rede, tornando o treinamento mais difícil e demorado, pois dobrou a dimensão dos vetores de entrada se comparados aos modelos que utilizavam somente dados dinâmicos ou estáticos.

Também foi criada uma rede neural de fusão dos dados exógenos e indicadores com os dados dependentes do ativo. Para tanto foram utilizados os modelos Base de cada uma das RNAs, contemplando a fusão de todas as variáveis disponíveis. Para este experimento foram utilizadas as variáveis dinâmicas formadas pelo cálculo da diferença relativa entre 2 dias. Os resultados obtidos não foram melhores do que aqueles obtidos com os dados dependentes em uma RNA e com os dados exógenos e indicadores em outra RNA.

Portanto, nenhuma das formas de fusão de características proporcionou resultados melhores na previsão de tendência. Com isso, pode-se afirmar que não existe uma complementariedade nas informações dos dados de entrada, que proporcione um desempenho melhor na previsão de tendência de movimentação do preço das ações analisadas.

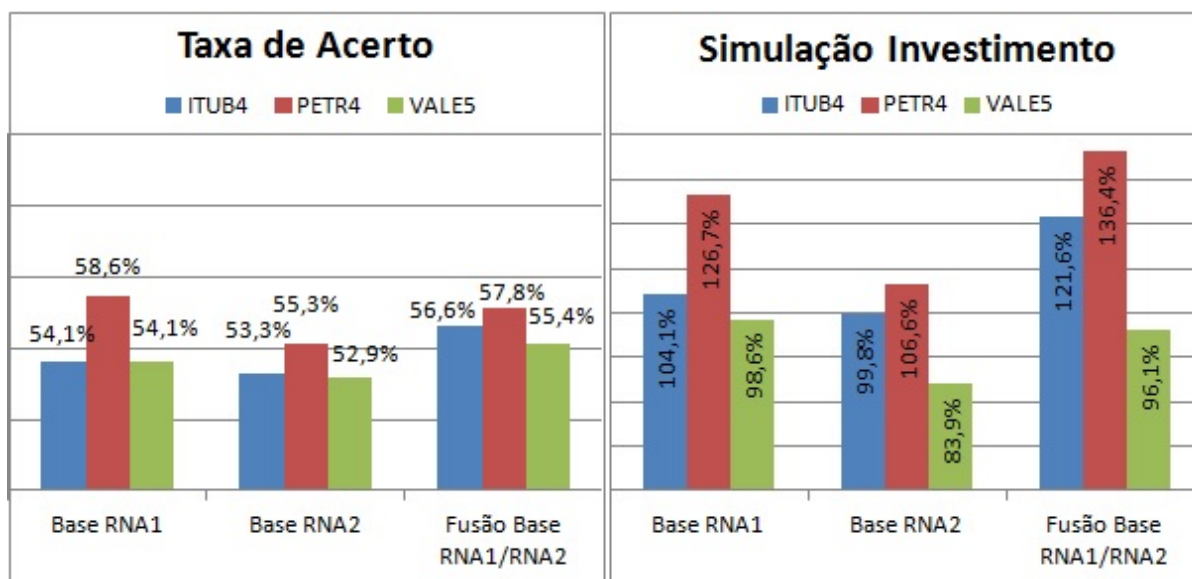
5.4.2 Fusão dos Sistemas Especialistas

De acordo com o princípio *dividir e conquistar*, uma tarefa complexa pode ser solucionada pela divisão em tarefas mais simples. Cada uma dessas sub-tarefas é direcionada para sistemas especialistas que geram soluções para elas. Após, esses sistemas são combinados para chegar a uma solução supostamente melhor do que aquela alcançável por qualquer especialista atuando isoladamente (HAYKIN, 2001). Essa combinação de sistemas especialistas também é conhecida como máquina de comitê.

Um dos experimentos de fusão de sistemas especialistas realizado, consistiu na fusão dos modelos Base com dados dinâmicos da diferença de 2 dias. Neste caso, os resultados apresentados por uma das RNAs que foi combinada, a RNA com variáveis exógenas e indicadores, para a ação VALE5, apontavam um prejuízo de 16,1%. A fusão dos sistemas produziu resultados melhores, reduzindo o prejuízo apresentado pela RNA com variáveis exógenas e indicadores. E quando os resultados das RNAs já apresentavam lucro, como para a ação PETR4, a fusão dos especialistas possibilitou alcançar resultados ainda melhores. Essa melhoria alcançada pela fusão de especialistas é mostrada na [Figura 22](#).

Outro modelo de fusão que merece destaque é o formado pela fusão dos modelos Base com atraso, que utilizam dados dinâmicos calculados pela diferença de 3 dias. Este alcançou bons retornos financeiros na simulação de negociação das 3 ações analisadas, como mostra a [Figura 23](#), possibilitando a obtenção de lucro de 13,8%, nas ações ITUB4 e

Figura 22: Taxa de acerto e simulação de investimento da fusão dos modelos Base da RNA com variáveis dependentes (1) e da RNA com variáveis exógenas e indicadores (2), utilizando dados dinâmicos calculados pela diferença de 2 dias.



Fonte: desenvolvido pelo autor

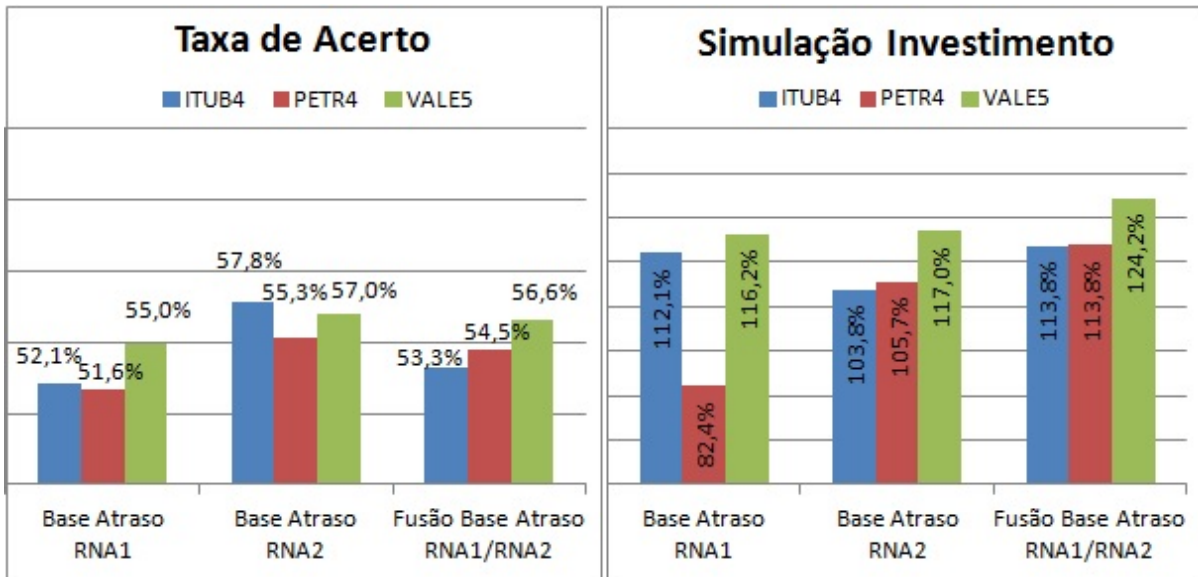
PETR4, e de 24,2% na ação VALE5. Destaca-se que, a fusão destes modelos não acarretou em um aumento da taxa de acerto, porém proporcionou maiores ganhos no investimento simulado para as 3 ações analisadas.

Foram realizados diversos outros experimentos de fusão de modelos das RNAs com variáveis dependentes e modelos das RNAs com variáveis exógenas e indicadores. A grande maioria dos experimentos de fusão possibilitou alcançar resultados melhores que aqueles alcançados pelas RNAs individualmente e, em situações adversas, amenizou possíveis prejuízos com uma RNA de baixo desempenho.

Após todos os experimentos realizados, este trabalho aponta como sistema predictor final, o sistema composto pela fusão dos modelos Base com atraso, com dados dinâmicos da diferença de 3 dias (Figura 23). Este foi o sistema escolhido pois possibilitou a obtenção de ganhos em todas as ações analisadas, sendo mais consistente que os outros sistemas.

Para que se possa ter uma melhor noção do desempenho do sistema predictor final, este foi comparado com a variação real do preço das ações no período analisado e com um sistema que gera tendências aleatórias. As ações analisadas tiveram uma desvalorização significativa no período, no qual a PETR4 teve queda de 29,6%, a VALE5 teve queda de 24,8% e a ITUB4 teve queda de 16,9%. Se o investidor negociasse as ações seguindo um sistema que gerasse tendências aleatórias, teria um prejuízo um pouco menor, de 14,7% na ação PETR4, de 10,1% na ação VALE5 e de 8,1% na ação ITUB4. Estes valores são um

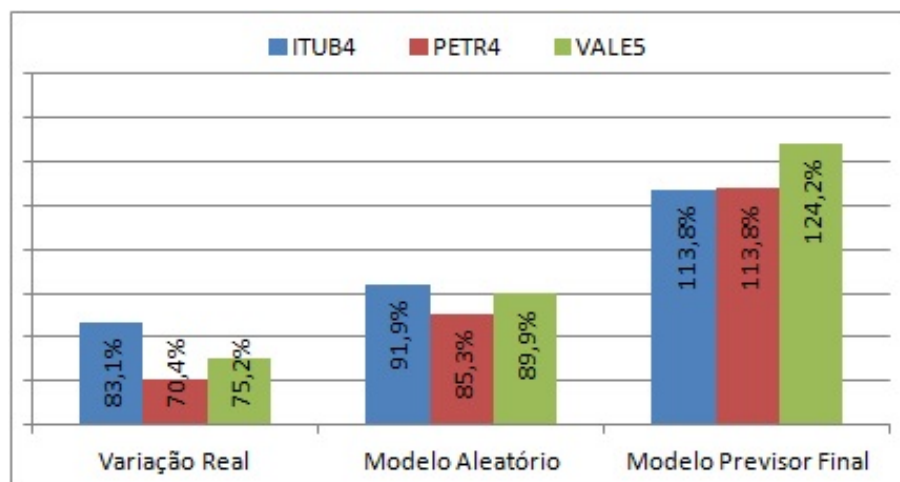
Figura 23: Taxa de acerto e simulação de investimento da fusão dos modelos Base com Atraso da RNA com variáveis dependentes (1) e da RNA com variáveis exógenas e indicadores (2), utilizando dados dinâmicos da diferença de 3 dias.



Fonte: desenvolvido pelo autor

resultado médio, obtido após 100 repetições da simulação de investimento com o sistema aleatório. Já se o investidor tivesse investido seguindo as tendências geradas pelo modelo de previsão final, este obteria lucros maiores que 13% nas 3 ações. Estas informações são mostradas na [Figura 24](#).

Figura 24: Variação real das ações no período analisado, simulação de investimento com tendências aleatórias e com tendências geradas pelo modelo predictor final.



Fonte: desenvolvido pelo autor

6 CONSIDERAÇÕES FINAIS

A previsão do comportamento futuro das ações é essencial para que o investidor possa obter ganhos financeiros no mercado acionário. Tradicionalmente, a análise de ações é realizada de duas formas: fundamentalista e técnica. A análise fundamentalista tem um caráter de projeção de médio a longo prazo, está baseada nos fundamentos da empresa e nas perspectivas dela dentro de um contexto econômico. Já a análise técnica, tem como premissa que o comportamento dos preços se repete ao longo do tempo e, por este motivo, entende-se que uma análise do histórico do ativo em questão é suficiente para prever o seu futuro.

Ambas as formas de análise possuem pontos negativos. A análise fundamentalista exige que o analista tenha um grande conhecimento contábil e do contexto econômico para realizar uma projeção de futuro. Já na análise técnica a investigação dos dados históricos, além de ser demorada, pode ser complexa, uma vez que a distribuição dos dados pode dificultar a identificação de padrões do movimento dos preços. Para facilitar e automatizar esses processos de reconhecimento de padrões podem ser utilizados sistemas especialistas computacionais.

Através de um processo de aprendizagem, os sistemas especialistas adquirem os conhecimentos inerentes a uma aplicação e os utilizam de forma genérica na solução de problemas similares. Quando a tarefa é dita complexa, ela pode ser dividida entre vários sistemas especialistas para, após a execução de todos estes sistemas, combinar os resultados obtidos por cada um deles individualmente.

O presente trabalho teve como objetivo desenvolver uma aplicação de previsão da tendência de movimentação dos preços de ações baseada na combinação de sistemas especialistas. Normalmente, para este tipo de sistema especialista, RNAs são utilizadas. A RNA é uma técnica de inteligência computacional que tenta simular o comportamento do cérebro humano. Seu aprendizado é adquirido através de experiências, onde para cada valor de entrada fornecido para a RNA o valor de saída desejado é conhecido. Quando ocorre algum erro a RNA aprende e se adapta a estas novas informações.

Os dados de entrada fornecidos para a RNA devem carregar consigo características que auxiliem a identificação de padrões. As RNAs realizam a extração destas características e, a partir delas, classificam os dados de entrada conforme os padrões pré-definidos. A partir da análise de trabalhos correlatos foram selecionadas as variáveis de entrada do sistema previsor. Elas podem ser divididas em três grupos:

- Variáveis dependentes: são dados diretamente relacionados com a ação, correspondem a um resumo diário histórico da comercialização daquela ação em bolsa;

- Variáveis exógenas: são independentes do ativo a ser previsto, mas podem ter alguma influência na formação de seu preço ou representar informações econômicas relevantes para o mercado de ações;
- Indicadores de análise técnica: são dados gerados a partir da série da cotação de fechamento, que formam indicadores de análise técnica e podem fornecer informações relacionadas com a tendência de movimentação do preço para a RNA.

O sistema predictor modelado foi composto por três RNAs. Cada uma destas RNAs foi treinada individualmente, de modo que as duas primeiras gerem as informações que são combinadas em uma terceira RNA, esta de fusão, visando melhorar a eficiência de previsão. Neste trabalho foram analisadas 3 ações dentre as de maior liquidez do mercado acionário brasileiro: ITUB4, PETR4 e VALE5.

Considerando o uso desta aplicação no mundo real, ela seria utilizada minutos antes do fechamento diário do pregão. Ela gera uma tendência de alta ou baixa nos preços do dia seguinte. Dessa forma, de acordo com a previsão do sistema, o investidor tem a possibilidade de negociar suas ações antes do fechamento do pregão do dia atual.

Para analisar o desempenho do sistema predictor desenvolvido, foram utilizadas duas métricas: a taxa de acerto de tendência e uma simulação de investimento. As taxas de acerto de tendência obtidas com os sistemas não ultrapassaram a marca dos 59%. Do ponto de vista da área de reconhecimento de padrões, 59% não é um resultado muito significativo. Por isso é importante analisar o sistema do ponto de vista financeiro. Como forma de avaliação financeira, foi aplicada sobre o sistema predictor uma simulação de investimento, realizada sobre os dados do período de março de 2013 à fevereiro de 2014 (1 ano).

Nesse período as ações analisadas tiveram uma desvalorização significativa, no qual as 3 ações tiveram queda superior a 16%. Mesmo com essa queda importante no preço das ações, o sistema predictor final, possibilitaria a um investidor, se este seguisse as tendências de alta/baixa geradas pelo sistema, obter lucros superiores a 13% em cada uma das 3 ações. O sistema final, apesar de não alcançar uma taxa de acerto elevada, se mostrou bem mais eficiente que um sistema que gera tendências aleatórias, que levaria o investidor a ter prejuízo de mais de 8% nas ações durante o período.

Para explicitar o quão bom foi o desempenho obtido pelo sistema predictor foi calculada a diferença absoluta entre os percentuais obtidos com o sistema predictor e os percentuais de variação real da ação. Na ação ITUB4 a diferença absoluta entre o sistema predictor e a variação real da ação foi de 30,7%. Na ação PETR4 a diferença absoluta foi de 43,4% e na ação VALE5 a diferença foi ainda maior, atingindo 49%. Portanto, em todas as ações o sistema predictor alcançou percentuais bastante elevados se comparados com a variação real da ação no período. Considerando que, em momentos de alta é mais

fácil obter lucros no mercado de ações e que o sistema previsor alcançou percentuais bem superiores, conclui-se que este sistema previsor tem capacidade de gerar lucros ainda maiores em momentos mais favoráveis no mercado de ações.

A partir do presente trabalho vislumbram-se algumas possibilidades de continuidade. Uma delas é tornar o protótipo desenvolvido um sistema com interface gráfica que possibilite ao investidor utilizá-lo com facilidade, no qual os dados diários da comercialização das ações sejam automaticamente capturados pelo sistema previsor e este gere a tendência do próximo dia. Outra possibilidade de continuidade é testar este sistema com outras ações do mercado brasileiro, assim como em diferentes períodos, para verificar sua capacidade de gerar lucros.

A nível de pesquisa, poderiam ser agregadas novas informações ao sistema previsor, como outros indicadores de análise técnica, indicadores de análise fundamentalista e/ou notícias relevantes ao mercado acionário, de modo que estas informações poderiam auxiliar no processo de previsão de tendências. Outra pesquisa poderia ser dirigida para responder como a tendência é afetada por cada uma das características de entrada utilizadas neste trabalho.

REFERÊNCIAS

- ABRAHAM, A.; PHILIP, N. S.; SARATCHANDRAN, P. Modeling chaotic behavior of stock indices using intelligent paradigms. **arXiv preprint cs/0405018**, 2004.
- BAHRAMY, F.; CRONE, S. F. Forecasting foreign exchange rates using support vector regression. In: IEEE. **Computational Intelligence for Financial Engineering & Economics (CIFEr)**, 2013 IEEE Conference on. [S.l.], 2013. p. 34–41.
- BM&FBOVESPA. **Como Investir em Ações**. [S.l.: s.n.], 2011. Disponível em: <<http://lojavirtual.bmf.com.br/LojaIE/Portal/Pages/pdf/ComoInvestir.pdf>>. Acesso em 22 de Agosto de 2014.
- BM&FBOVESPA. **Site Oficial da BM&FBOVESPA**. 2014. Acesso em 22 de Outubro de 2014. Disponível em: <<http://www.bmfbovespa.com.br/>>.
- BORGES, M. Algoritmo genético para otimização de arquiteturas de redes mlp. 2012.
- Bússola do Investidor. **Bússola do Investidor**. 2014. Acesso em 15 de Novembro de 2014. Disponível em: <<http://www.bussoladoinvestidor.com.br/>>.
- CAMARGOS, F. L. Lógica nebulosa: uma abordagem filosófica e aplicada. **Florianópolis: UFSC**, 2002.
- CASAGRANDE NETO, H.; SOUSA, L. A.; ROSSI, M. C. **Abertura de Capital de Empresas no Brasil: Um enfoque prático**. 4. ed. São Paulo: Atlas, 2010.
- CASTANHEIRA, N. P. **Métodos quantitativos**. [S.l.]: Editora Ibplex, 2008.
- CHARKHA, P. R. Stock price prediction and trend prediction using neural networks. In: IEEE. **Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on**. [S.l.], 2008. p. 592–594.
- CORREIA, L. F.; AMARAL, H. F.; BRESSAN, A. A. O efeito da liquidez sobre a rentabilidade de mercado das ações negociadas no mercado acionário brasileiro. **Revista Base (Administração e Contabilidade) da UNISINOS**, Universidade do Vale do Rio dos Sinos, v. 5, n. 2, p. 109–119, 2008.
- DOUGHERTY, G. **Pattern Recognition and Classification**. Camarillo, CA, USA: Springer, 2013.
- FARIA, E. L. d.; ALBUQUERQUE, M. P.; ALFONSO, J. L. G.; ALBUQUERQUE, M.; CAVALCANTE, J. Previsão do mercado de ações brasileiro utilizando redes neurais artificiais. **CBPF-Centro Brasileiro de Pesquisas Físicas-Novas Técnicas e PUC-RJ, Rio de Janeiro**, 2008.
- FORTUNA, E. **Mercado Financeiro: Produtos e serviços**. 18. ed. Rio de Janeiro: Qualitymark, 2011.
- FREEMAN, J. A.; SKAPURA, D. M. **Neural Networks: Algorithms, applications, and programming techniques**. [S.l.]: Citeseer, 1991.

- FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**. 2. ed. San Diego, USA: Academic Press, 1990.
- HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre: Bookman, 2001.
- HE, H.; GARCIA, E. A. Learning from imbalanced data. **Knowledge and Data Engineering, IEEE Transactions on**, IEEE, v. 21, n. 9, p. 1263–1284, 2009.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural networks**, Elsevier, v. 2, n. 5, p. 359–366, 1989.
- KERR, R. B. **Mercado de Financeiros e de Capitais**. São Paulo: Pearson Prentice Hall, 2011.
- KHAN, A. U.; BANDOPADHYAYA, T.; SHARMA, S. Comparisons of stock rates prediction accuracy using different technical indicators with backpropagation neural network and genetic algorithm based backpropagation neural network. In: IEEE. **Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on**. [S.l.], 2008. p. 575–580.
- LI, R.-J.; XIONG, Z.-B. Forecasting stock market with fuzzy neural networks. In: IEEE. **Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on**. [S.l.], 2005. v. 6, p. 3475–3479.
- LI, Y.; MA, W. Applications of artificial neural networks in financial economics: a survey. In: IEEE. **Computational Intelligence and Design (ISCID), 2010 International Symposium on**. [S.l.], 2010. v. 1, p. 211–214.
- LIMA, Á. H. N. d. Seleção de características de dados utilizando redes neurais artificiais. 2015.
- LUGER, G. F. **Inteligência Artificial**. São Paulo: Pearson Education do Brasil, 2013.
- MACEDO, J.; CAMARGO, L. O.; OLIVEIRA, H. Brandao de; SILVA, L. da; SALGADO, R. M. An intelligent decision support system to investment in the stock market. **Latin America Transactions, IEEE (Revista IEEE America Latina)**, IEEE, v. 11, n. 2, p. 812–819, 2013.
- MACIEL, L. S.; BALLINI, R. Design a neural network for time series financial forecasting: Accuracy and robustness analisys. 2009.
- MAGER, J.; PAASCHE, U.; SICK, B. Forecasting financial time series with support vector machines based on dynamic kernels. In: IEEE. **Soft Computing in Industrial Applications, 2008. SMCia'08. IEEE Conference on**. [S.l.], 2008. p. 252–257.
- MARQUES, F. C.; GOMES, R. M. Análise de séries temporais aplicadas ao mercado financeiro com o uso de algoritmos genéticos e lógica nebulosa. **VII Encontro Nacional de Inteligência Artificial. Bento Gonçalves: Sociedade Brasileira de Computação**, p. 749–768, 2009.
- MORETTIN, P. A.; TOLOI, C. M. **Séries Temporais**. 2. ed. São Paulo: Atual, 1987.
- MOTTA, V. T. da. **Bioestatística**. 2. ed. Caxias do Sul, RS: EDUCS, 2006. 190 p.

- MURPHY, J. J. **Technical analysis of the financial markets: A comprehensive guide to trading methods and applications**. [S.l.]: Penguin, 1999.
- NAEINI, M. P.; TAREMIAN, H.; HASHEMI, H. B. Stock market value prediction using neural networks. In: IEEE. **Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on**. [S.l.], 2010. p. 132–136.
- NETO, M. Christovam de A.; CAVALCANTI, D. da C. et al. **Previsão de séries temporais usando séries exógenas e combinação de redes neurais aplicada ao mercado financeiro**. Dissertação (Mestrado), 2008.
- PEREIRA, C. L. **Mercado de Capitais**. Curitiba: [s.n.], 2013. Livro Eletrônico.
- RIBEIRO, C.; GOLDSCHMIDT, R.; CHOREN, R. Métodos para previsão de séries temporais e suas tendências de desenvolvimento. **Monografia, Instituto Militar de Engenharia, Brasil**, 2009.
- SACHETIM, H. M. **Análise Técnica: Estudo da Confiabilidade dos Principais Indicadores de Análise Técnica, aplicados as ações mais negociadas na Bovespa no período de 1995 a 2005**. Dissertação (Mestrado) — Universidade Federal do Paraná, 2006.
- SAFFI, P. A. Análise técnica: sorte ou realidade? **Revista Brasileira de Economia, SciELO Brasil**, v. 57, n. 4, p. 953–974, 2003.
- SOONG, F. K.; ROSENBERG, A. E. On the use of instantaneous and transitional spectral information in speaker recognition. **Acoustics, Speech and Signal Processing, IEEE Transactions on**, IEEE, v. 36, n. 6, p. 871–879, 1988.
- STOCK, J. H.; WATSON, M. W. **Econometria**. São Paulo: Addison Wesley, 2004.
- TAY, F. E.; CAO, L. Application of support vector machines in financial time series forecasting. **Omega**, Elsevier, v. 29, n. 4, p. 309–317, 2001.
- THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. 2. ed. San Diego: Elsevier, 2003.
- VIDOTTO, R. S.; MIGLIATO, A. L. T.; ZAMBON, A. C. O moving average convergence moving average convergence moving average convergence-divergence divergence divergence como ferramenta para a decisão de investimentos no mercado de ações a decisão de investimentos no mercado de ações. **SciELO Brasil**, 2009.
- VUI, C. S.; SOON, G. K.; ON, C. K.; ALFRED, R.; ANTHONY, P. A review of stock market prediction with artificial neural network (ann). In: IEEE. **Control System, Computing and Engineering (ICCSCE), 2013 IEEE International Conference on**. [S.l.], 2013. p. 477–482.
- WANG, Z.; BOVIK, A. C. Mean squared error: love it or leave it? a new look at signal fidelity measures. **Signal Processing Magazine, IEEE, IEEE**, v. 26, n. 1, p. 98–117, 2009.
- WEBB, A. R. **Statistical Pattern Recognition**. 2. ed. Chichester: John Wiley & Sons Ltd, 2002.

WU, J.-Y.; LU, C.-J. Computational intelligence approaches for stock price forecasting. In: IEEE. **Computer, Consumer and Control (IS3C), 2012 International Symposium on**. [S.l.], 2012. p. 52–55.

WUERGES, A. F. E.; BORBA, J. A. Redes neurais, lógica nebulosa e algoritmos genéticos: aplicações e possibilidades em finanças e contabilidade. **Journal of Information Systems and Technology Management**, SciELO Brasil, v. 7, n. 1, p. 163–182, 2010.

YILDIZ, B.; YALAMA, A.; COSKUN, M. Forecasting the istanbul stock exchange national 100 index using an artificial neural network. **World Academy of Science, Engineering and Technology**, Citeseer, v. 46, p. 36–39, 2008.

YOO, P. D.; KIM, M. H.; JAN, T. Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. In: IEEE. **Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on**. [S.l.], 2005. v. 2, p. 835–841.

ZHANG, D.; JIANG, Q.; LI, X. Application of neural networks in financial data mining. **International Journal of Computational Intelligence**, v. 1, n. 2, p. 116–119, 2004.