

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

Valter Pereira de Carvalho

PREVISÃO DE SÉRIES TEMPORAIS NO MERCADO
FINANCEIRO DE AÇÕES COM O USO DE REDE NEURAL
ARTIFICIAL

São Paulo
2018

**UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO**

Valter Pereira de Carvalho

**PREVISÃO DE SÉRIES TEMPORAIS NO MERCADO
FINANCEIRO DE AÇÕES COM O USO DE REDE NEURAL
ARTIFICIAL**

Dissertação para o Programa de Pós-Graduação
da Universidade Presbiteriana Mackenzie como
requisito parcial para a obtenção do título de
Mestre em Engenharia Elétrica e Computação.

Orientador: Prof.Dr. Leandro Augusto da Silva

São Paulo
2018

C331p

Carvalho, Valter Pereira de
Previsão de séries temporais no mercado financeiro de ações com o
uso de rede neural artificial. / Valter Pereira de Carvalho -
São Paulo, 2018.

70 f. ; 30 cm.

Dissertação (Mestrado em Engenharia Elétrica e Computação) -
Universidade Presbiteriana Mackenzie - São Paulo, 2018.
Orientador: Leandro Augusto da Silva.
Bibliografia: f. 54-59.

1. Rede neural. 2. RNA. 3. MLP. 4. Algoritmos. 5. BackPropagation.
 6. Random Walk. 7. BOVESPA. 8. Mercados Eficientes. 9. Previsão.
 10. Mercado financeiro.
- I. Silva, Leandro Augusto, orientador. II. Título.

CDD 004.015192

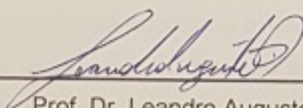
VALTER PEREIRA DE CARVALHO

PREVISÃO DE SÉRIES TEMPORAIS NO MERCADO FINANCEIRO DE
AÇÕES COM O USO DE REDE NEURAL ARTIFICIAL

Dissertação de Mestrado apresentada
ao Programa de Pós-Graduação em
Engenharia Elétrica e Computação da
Universidade Presbiteriana Mackenzie,
como requisito parcial para a obtenção
do título de Mestre em Engenharia
Elétrica e Computação.

Aprovada em 03 de agosto de 2018.

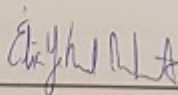
BANCA EXAMINADORA



Prof. Dr. Leandro Augusto da Silva
Orientador - Universidade Presbiteriana Mackenzie



Prof. Dr. Paulo Rogério Scarano
Universidade Presbiteriana Mackenzie



Prof.ª Dr.ª Elia Yathie Matsumoto
Universidade de São Paulo

Figura 1: Ficha da Banca

Dedico ao meu pai (in memorian) por me fazer amar e acreditar no alfabeto como única
forma de mudar a vida.

AGRADECIMENTOS

Ao meu orientador de mestrado Prof. Dr. Leandro Augusto da Silva pela paciência, confiança, amizade, fundamentalmente pela insistência em me manter vivo quando tudo parecia não mais acontecer nesse mestrado e mais ainda pela ajuda extra que foi grande sustentáculo para que eu chegasse até aqui.

Aos meus professores do Mestrado Profissional de Economia e Mercados, Prof. Dr. Paulo Rogério Scarano e Prof. Dr. Vladimir Fernandes Maciel, ambos forneceram bases sólidas seja no âmbito pedagógico como pessoal e que foram fundamentais para que esse trabalho existisse.

Aos colegas de turma, professores e funcionários do Programa de Mestrado em Engenharia Elétrica e Computação da Universidade Mackenzie pelo companheirismo de sempre.

Para a minha mãe Anésia Minelvina de Carvalho por todo o esforço que fez trabalhando duro para que eu sempre estudasse nas melhores escolas, no caso a Universidade Presbiteriana Mackenzie.

Para a minha irmã Rosangela Pereira de Carvalho pelas ajudas que sempre foram trazidas nos momentos difíceis e pelos incentivos para eu estudar cada vez mais.

Aos meus filhos Paula Leite de Carvalho e Walter Pereira de Carvalho Filho por todo o carinho que me dedicam e pelo incentivo para que eu nunca pare de estudar.

Por fim àquela que é o meu verdadeiro porto seguro emocional, espiritual e intelectual, minha esposa Marli Leite de Carvalho, que desde que nos conhecemos fez com que a minha vida se completasse em todos os sentidos e que sempre me incentiva a permanecer no meio acadêmico, seja como professor ou como aluno. Foram inúmeros dias de presença física e ausência ao mesmo tempo, mas sempre com seu apoio para que os nossos objetivos fossem atingidos.

RESUMO

Este trabalho propõe um estudo de previsão de séries temporais com o uso dos dados obtidos da BOVESPA (Bolsa de Valores de São Paulo) tomando-se por base os valores das ações no fechamento do pregão. Para a previsão será utilizada uma rede neural artificial (RNA) com arquitetura MLP (MultiLayer Perceptron). Será mostrado através desse estudo de previsão do mercado financeiro como a rede neural se comporta e como ela pode ser de grande valia para previsões com séries de dados temporais. A análise compreende a comparação entre a previsão e o preço de fechamento efetivo dentro de períodos estabelecidos. O trabalho faz um comparativo entre a rede MLP e a Hipótese de Random Walk. Ao final do trabalho conclui-se que a rede neural artificial utilizada para previsão de mercado acionário é capaz de mostrar resultados muito próximos da realidade, e que essa metodologia pode ser utilizada por investidores individuais e coletivos para compreenderem o comportamento das ações e se orientarem sobre as possíveis hipóteses de investimentos.

Palavras-chave: *rede neural, RNA, MLP, BOVESPA, Random Walk, previsão, mercado financeiro.*

ABSTRACT

This work proposes a study of the forecast of time series with the use of data obtained from BOVESPA the basis of the values of the shares at the closing of the trading session. For the forecast, an artificial neural network (RNA) with MLP (MultiLayer Perceptron) architecture will be used. It will be shown through this prediction study of the financial market how the neural network behaves and how it can be of great value for forecasts with time series data. The analysis comprises the comparison between the forecast and the effective closing price within established periods. The paper compares the MLP network with the Random Walk Hypothesis. At the end of the study it is concluded that the artificial neural network used for stock market forecasting is able to show results very close to reality, and that this methodology can be used by individual and collective investors to understand the behavior of the actions and to orient themselves on the possible investment hypotheses.

Keywords: *neural network, RNA, MLP, BOVESPA, Random Walk, forecasting, financial market.*

Lista de Figuras

1	Ficha da Banca	4
2	Neurônio Artificial - Fonte: Adaptado pelo autor de Nunes, 2016	24
3	Neurônio de McCulloch e Pitts - 1943 - Fonte: Adaptado pelo autor de McCulloch e Pitts (1943)	27
4	Rede Perceptron - Fonte: Adaptado pelo autor de HAYKIN, 2001	29
5	Rede MLP - Fonte: Adaptado pelo autor de HAYKIN, 2001	32
6	Fluxo de execução da metodologia experimental - Fonte: Autor	36
7	Visão parcial da série de dados de entrada na rede neural - Fonte: Autor	39
8	Distribuição quantitativa das amostras - Fonte: Autor	39
9	Cenário 1 - Série em nível - Fonte: Autor	41
10	Cenário 2 - Série em nível - Fonte: Autor	41
11	Resultado dos testes ADF sobre os dados da série em nível	42
12	Cenário 1 - Série da primeira diferença - Fonte: Autor	42
13	Cenário 2 - Série da primeira diferença - Fonte: Autor	43
14	Resultado dos testes ADF sobre os dados da série da primeira diferença	43
15	Quadro comparativo do Erro MSE para vários ensaios por cenário	44
16	Legenda	45
17	Experimentos selecionados para comparação com Random Walk	45
18	Erro x Taxa de aprendizagem para experimentos com 5 neurônios	45
19	Erro x Taxa de aprendizagem para experimentos com 10 neurônios	46
20	Cenário 1 - Experimento 8 - Previsão - MLP	46
21	Experimento 8 - Diferença previsto e observado - MLP	47
22	Cenário 1 - Experimento 8 - Dados da Previsão - MLP	47
23	Cenário 2 - Experimento 15 - Previsão - MLP	48
24	Cenário 2 - Experimento 15 - Diferença previsto e observado - MLP	48
25	Cenário 2 - Experimento 15 - Dados da Previsão	49
26	Cenário 1 - Experimento 8 - Previsão - RW (Random Walk)	50
27	Cenário 2 - Experimento 15 - Previsão - RW (Random Walk)	50
28	Quadro comparativo dos resultados entre a rede neural MLP com a hipótese de Random Walk	51

Sumário

1	INTRODUÇÃO	1
1.1	Contextualização	1
1.2	Descrição do Problema	3
1.3	Objetivo Geral	5
1.4	Objetivo Específico	5
1.5	Justificativa	5
1.5.1	Contribuições da pesquisa	7
2	REFERENCIAL TEÓRICO	7
2.1	Fundamentos sobre Mercado de Capital	8
2.1.1	Conceito de Ações	9
2.1.2	Tipos de Ações e sua classificação segundo a Lei das S.A.	10
2.1.3	Ações quanto a circulação	11
2.1.4	Os rendimentos e riscos das ações	11
2.1.5	Os tipos de mercado	13
2.1.6	O financiamento da empresa	14
2.1.7	A Bolsa de Valores	14
2.1.8	O que é um Derivativo de Crédito	15
2.2	Preparação de Séries Temporais	17
2.2.1	Conceito de Séries Temporais	17
2.2.2	Séries Temporais Estacionárias	19
2.2.3	Contornando o problema	20
2.2.4	Modelo Matemático	20
2.2.5	Análise Exploratória e Pré-Processamento dos dados	21
2.3	Fundamentos sobre Redes Neurais Artificiais	22
2.3.1	A Inteligência Artificial e as Redes Neurais Artificiais	22
2.3.2	Conceito de Redes Neurais Artificiais (RNA)	22
2.3.3	Estrutura das Redes Neurais Artificiais	23
2.3.4	Fatores de desempenho das RNAs	24
2.3.5	Tipos de Redes Neurais	26
2.3.5.1	Redes neurais de camada única – Perceptrons	28
2.3.5.2	Redes neurais lineares	29

2.3.5.3	Redes neurais multicamada – Multilayer Perceptrons . . .	30
2.3.6	Tipos de Algoritimos	32
2.3.6.1	Algoritmo Backpropagation	32
2.3.6.2	Algoritmo de Random Walk	34
2.3.7	Modelos Utilizados	35
3	METODOLOGIA	35
3.1	Procedimentos Preparatórios	36
3.2	Seleção dos dados das ações PETR4	37
3.3	Análise Estatística	37
3.4	Separação dos dados para treinamentos	38
3.5	Execução das previsões	39
3.6	Avaliação dos resultados	40
4	EXPERIMENTOS E RESULTADOS	40
5	CONCLUSÕES	53
	REFERÊNCIAS BIBLIOGRÁFICAS	59

1 INTRODUÇÃO

1.1 Contextualização

A Econometria é um ramo da Economia, que pode ser entendida como um conjunto de ferramentas estatísticas que utiliza variáveis econômicas, como entradas de modelos matemáticos, para interpretar os fenômenos econômicos. Na econometria busca-se em última análise descrever os fenômenos econômicos matematicamente e com isso poder prever as ocorrências futuras, sejam essas de ganhos, perdas ou equilíbrio(WOOLDRIDGE, 2016).

Assim, no Mercado Financeiro que opera a compra e venda de títulos de créditos (ações, debentures e outros), a grande preocupação de seus operadores, no caso as corretoras e corretores, é tentar acertar as tendências de determinada ação para assim evitar perdas, ou prever o lucro para assim executar o investimento e realizar o lucro no resultado. Para isso, hoje se utilizam de várias planilhas e ou softwares específicos que fazem uma série de análises sobre os dados históricos, buscando uma tendência do mercado.

É importante ressaltar que o comportamento do Mercado Financeiro é bastante sensível a uma série de fenômenos que interferem diretamente nos resultados dos valores das ações, tais como assimetria da informação, fatos políticos, fatos sociais e empresariais, mudança de expectativa de determinado investidor que detenha grande capacidade de investimento, fatos climáticos e outros.Tais fenômenos justificam a dificuldade em uma modelagem analítica que torne a previsão um valor confiável para tomada de decisões com o menor risco possível, remetendo assim aos estudos da eficiência do mercado (WOOLDRIDGE, 2016).

A eficiência de mercado é um tema muito avaliado relevante em Mercado Financeiro, inicialmente estudado, por Paul Samuelson, e a posteriori analisado por Eugene Fama. Nos estudos de finanças, um mercado de capitais eficiente é quando se pode incorporar na determinação do preço toda informação disponível do ativo transacionado do mercado.(FAMA, 1970).

Por outro lado, a política monetária adotada por um país interfere fortemente no desempenho do seu mercado de capital. As variações na taxa de juros podem provocar um incentivo à migração dos investimentos de mercado de capitais para outros investimentos

como títulos do governo em um cenário de alta da taxa de juros ou em uma migração para o mercado de capitais no cenário oposto, ou seja, em caso de queda da taxa. Taxas de juros elevadas via de regra geral estimulam os investidores a se desfazerem de sua posição no mercado de renda variável e a investirem em renda fixa. Avaliando-se as razões para esse comportamento pode-se citar duas principais: a primeira é a taxa de juros determinada pelo Banco Central que impacta diretamente a capacidade das empresas de contrair empréstimos, e a segunda o apetite dos investidores(CASTRO, 2003).

Há uma considerável literatura americana sobre testes em modelos de passeios aleatórios ou random walk, com destaque inicial para o artigo: Random Walks in Stock Market Prices, de Eugene F. Fama que aborda os preços das ações num Mercado Eficiente (ou pelo menos Semi-Eficiente), assim como também The Random-Walk Theory: An Empirical Test, artigo de James C. Van Horne e George G.C. Parker, dentre outros.

A aplicação de técnicas de predição em séries temporais financeiras é um assunto de constante e elevado interesse na comunidade científica, tanto para investidores quanto para pesquisadores, trata-se de uma área que traz grandes desafios em relação à complexidade dos problemas e que pode, eventualmente, gerar altos lucros financeiros para empresas envolvidas com análise e gerenciamento de ativos e com negociação automática em bolsas de valores(SCOTT, 1987a).

Na resolução do problema por meio de técnicas de predição, procura-se estimar o modelo que melhor aproxima a relação entre as entradas e a saída desejada (denominada "alvo"). Na sequência o modelo é extrapolado para estimar o valor futuro da série. Não é incomum pesquisas que realizam a predição para mais de um período (t), ou seja, para $t + 2; \dots; t + N$, onde N é o número de períodos adicionais, todavia o nível de erro acumulado encontrado neste tipo de modelo é normalmente muito alto e proibitivo na utilização.

O trabalho aqui apresentado fará uso de um conceito da computação e Inteligência Artificial em gênero, chamada de RNA (Rede Neural Artificial). As RNAs vem sendo utilizadas para várias aplicações que envolvem dados em séries temporais que é o caso do movimento de ações nas bolsas de valores. Ainda, as RNAs têm se mostrado bastante eficientes nos resultados com grandes massas de dados. Os trabalhos acadêmicos de (CARTACHO, 2001), (CASTRO, 2003) e (FREITAS, 2001) sobre o uso de RNAs para previsão de valores futuros de ativos financeiros são mostras que motivaram essa aplicação

e desenvolvimento.

A RNA é um método computacional para analisar sistemas que tenham comportamentos não lineares. Com inspiração no funcionamento do cérebro humano, a RNA possui um processo de aprendizagem (análise dos dados), generalização (validação dos dados), teste e por último a fase de previsão dos resultados. A saída final da rede é a representação do valor da função de ativação. Os neurônios tem a capacidade de aprendizagem, ajustando os pesos quando necessário. É essa aprendizagem que faz das redes neurais uma importante ferramenta para análise de dados ruidosos, que aqui são as oscilações dos preços das ações na bolsa de valores(SILVA, 2016b) e (HAYKIN, 1999).

O bom desempenho das redes neurais vem da capacidade de lidar com dados com muitas oscilações e perturbações periódicas e não periódicas. Essas características tornam as RNAs uma importante ferramenta para previsão de séries temporais, não apenas no mercado financeiro, mas com uso potencial na avaliação de créditos, análise de riscos, em diagnósticos médicos, reconhecimento de padrões visuais, sistemas de segurança, robótica, apenas para citar algumas aplicações dentre tantas possíveis(RUSSEL STUART J.; NORVIG, 2003).

1.2 Descrição do Problema

O problema tratado neste trabalho, com base nos estudos sobre Mercados Eficientes, é objetivamente comparar a aplicação de redes neurais artificiais com os processos utilizados na econometria para a previsão do valor futuro de uma ação negociada em bolsa de valores.

Analisando os comportamentos, dia a dia, dos valores registrados nas cotações das diversas ações comercializadas, denota-se um caráter aleatório nos resultados, mas há controvérsias. Para REFENES (1993, p.5), existem “fenômenos que parecem aleatórios, sem aparente periodicidade, embora recorrentes em um sentido geral.” Assim os eventos que alteram os preços no mercado, embora parecerem desconexos num sentido restrito, possuem por vezes características com eventos passados. Dessa forma essa não-aleatoriedade dos preços faz com que os mercados sejam previsíveis. A relação de dependência não-linear é uma forte justificativa para utilização de modelos econométricos e redes neurais artificiais.

Fazer previsões no sistema financeiro é uma tarefa complexa, pois os eventos não são regulares, são incertos e os dados por vezes são fragmentados em demasia(FREITAS, 2001). Nesse contexto grande parte dos modelos clássicos para predição utiliza apenas equações lineares, deduzidas e especificadas pelo usuário. Em modelos econômicos, pode-se levar muito tempo para a obtenção de resultados significativos, todavia, com a evolução e o avanço das tecnologias de informação e dos computadores, tornou-se possível desenvolver métodos mais apurados e as previsões.

A aplicação das RNAs busca tentar diminuir as dificuldades dos sistemas lineares e melhorar a qualidade dos resultados. As RNAs são capazes de processar grande quantidade de dados e identificar modelos e tendências, facilitando a tomada de decisões. Muitos bancos, corretoras de valores e empresas em geral estão usando as redes neurais para identificar relações e tendências complexas, cujo objetivo final é diminuir suas perdas, estabelecer análises sobre a volatilidade, buscar enfim melhorar o desempenho e resultados financeiros, detalhar e avaliar mais profundamente os riscos de créditos e de mercado e outras ações(REFENES, 1993).

Existem diversas teorias sobre análise e previsão do mercado acionário e todas tem objetivos claros na redução de riscos e otimização dos lucros e entre elas o modelo mais utilizado é o chamado *Capital Asset Pricing Model*, o CAPM (Em língua portuguesa: Modelo de Precificação de Ativos Financeiros (MPAF)). Esse modelo foi criado por Jack Treynor, William Forsyth Sharpe, John Lintner e Jan Mossin, independentemente, baseado no trabalho de Harry Markowitz sobre diversificação e teoria moderna de portfólio. Em 1990, Sharpe foi o vencedor do Prêmio Nobel de Economia, com Markowitz e Merton Miller, devido a sua contribuição ao campo de finanças.

O modelo CAPM é utilizado em finanças para determinar a taxa de retorno teórica apropriada de um determinado ativo em relação a uma carteira de mercado perfeitamente diversificada(FAMA, 1970). Ele leva em consideração a sensibilidade do ativo ao risco não-diversificável (também conhecido como risco sistêmico ou risco de mercado), representado pela variável conhecida como índice beta ou coeficiente beta, assim como o retorno esperado do mercado e o retorno esperado de um ativo teoricamente livre de riscos.

O CAPM estabelece uma relação linear positiva entre o ganho esperado de uma carteira de ativos diversificado num mercado dito eficiente. A hipótese dos mercados eficientes

proposta por (FAMA, 1970) afirma que o mercado é eficiente na alocação dos recursos, onde o preço dos ativos sempre reflete inteiramente todas as informações disponíveis, e todos os investidores dispõem dessas informações para tomada de decisão e não há informações assimétricas ou privilegiadas.

Após inúmeros testes de vários especialistas sobre a eficiência do CAPM a hipótese dos mercados eficientes dificilmente é mantida. A limitação desse modelo está na relação linear entre retorno e o beta, que é o índice utilizado para medir o risco não diversificável, pois não mostra com precisão os diferentes retornos entre os ativos. Dessa forma a perspectiva da limitação de modelos lineares que os modelos baseados em redes neurais se tornaram mais eficientes e essa é a hipótese fundamental desse trabalho, mostrar essa eficiência (BRITO NEY O.; MANAZES, 1981).

1.3 Objetivo Geral

O objetivo geral desse trabalho foi fazer um comparativo entre a Hipótese de Random Walk e a Rede Neural Artificial MLP (MultiLayer Perceptron) para a previsão futura dos preços das ações.

1.4 Objetivo Específico

Como objetivos específicos os seguintes pontos:

- a.) Definir os elementos estatísticos usados no algoritmo para treinamento da rede neural
- b.) Executar os refinamentos necessários apurando a devida acurácia da rede
- c.) Execução da previsão com base em intervalos já ocorridos para a verificação da aderência do resultado gerado pela rede e o resultado real.
- d.) Utilizar os mesmos dados na previsão através de Random Walk

1.5 Justificativa

A análise dos trabalhos de (CARTACHO, 2001), (CASTRO, 2003) e (FREITAS, 2001) mostra resultados importantes do uso das redes neurais para previsões do mercado acionário, esse trabalho se justifica na medida em que o modelo a ser apresentado pode atingir um grau de previsão diferenciado em relação aos modelos clássicos de séries temporais, cujas relações lineares limitam em demasia sua usabilidade na prática cotidiana.

A capacidade de modelagem de valores com relações não lineares, como é o caso do mercado de ações, justifica o uso de redes neurais artificiais, que apresentam níveis de erros considerados baixos se comparado com os modelos tradicionais de regressão lineares da econometria tradicional. Além disso, as redes neurais artificiais permitem acompanhar alterações contínuas e bruscas nas oscilações do mercado financeiro, o que não pode ser medido por um modelo matemático linear e pré-definido (RESENDE, 2016). Ainda é importante observar o que ocorre com a massa de dados obtida do mercado financeiro quando analisada pelos meios lineares, esses não são capazes de explicar as mudanças bruscas resultados da ocorrência de eventos aleatórios e que tem influência sobre os dados e consequentemente nos resultados. O mercado financeiro é extremamente sensível a qualquer mudança que ocorra no cenário (RESENDE, 2016) e (BARROS, 2018).

No atual estágio do mercado financeiro, com as economias interligadas internacionalmente e com os vários tipos de ativos sendo negociados todos os dias, é importante e indispensável a criação de novos instrumentos computacionais que otimizem os métodos tradicionais de predição, gerando resultados que possam ser de uso prático e em tempo real (RESENDE, 2016) e (BARROS, 2018).

Atualmente, observa-se um grande crescimento do uso das redes neurais artificiais para auxílio na tomada de decisões e, em boa parte deles utilizando, dados em séries temporais.

O uso de RNAs vem auxiliando diversas áreas do conhecimento, na medicina, por exemplo, no cruzamento de dados de pacientes com quadros clínicos semelhantes, onde com isso os médicos fazem prescrições mais adequadas; no sistema financeiro é utilizada para conter fraudes no uso de cartão de crédito, onde grandes bases de dados, denominadas data warehouse, são utilizadas em conjunto com redes neurais dando respostas rápidas a

eventuais alterações no perfil de consumo do cliente. Também a rede neural artificial é utilizada para aplicações climáticas utilizadas no setor agronegócio(FREITAS, 2001).

Portanto, esses resultados obtidos das redes neurais artificiais em suas várias aplicações e setores do conhecimento justificam a sua utilização nesse trabalho para a predição no mercado financeiro com o uso de séries temporais.

1.5.1 Contribuições da pesquisa

Os principais ganho percebido nesta pesquisa compreende: a utilização de um modelo adaptativo que se renova, por meio do janelamento e o reprocessamento das séries, a cada período de aplicação; e a simulação operacional muito próxima da realidade do mercado (com baixo índice de erro entre o real e o previsto), dada a operação do ativo real (valor da ação) e a implementação de restrições e contabilidade de custos transacionais.

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

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

Esse trabalho busca experimentar algo que a tecnologia computacional oferece e que vem sendo difundida pela ciência de dados, aplicável nas mais diversas áreas do conhecimento humano e da sociedade. Não há aqui a pretensão de se contrapor às grandes obras que construíram a ciência econômica e que foram laureadas com prêmios máximos da ciência, mas apenas de fornecer uma alternativa para mitigação da ansiedade humana em ser vitorioso no mercado financeiro.

2 REFERENCIAL TEÓRICO

O referencial teórico traz conceitos de economia sobre o tema abordado que é o mercado de ações, contido no item Fundamentos sobre o Mercado de Capital e os conceitos

de engenharia sobre a predição com o uso de redes neurais, contido no item Fundamentos sobre Redes Neurais Artificiais.

2.1 Fundamentos sobre Mercado de Capital

O mercado financeiro sempre causou expectativas na sociedade e estimulou a cobiça em seus operadores e investidores. Dessa feita tornou-se tarefa imprescindível analisar as cotações e desenvolver ferramentas de previsão dos preços futuros para as operações realizadas(LANARI, 2000).

O índice Dow Jones, criado em 1884 por Charles Dow, analisava os preços de 11 empresas, no papel quadriculado, sem qualquer auxílio de ferramentas automatizadas. No Brasil, a Bolsa de Valores do Rio de Janeiro surge em 1845 e a BOVESPA em 1890, época em que as cotações eram registradas com giz em um quadro negro. O IBOVESPA (índice da bolsa de valores de São Paulo) objeto desse trabalho foi criado em 1968(BONOMO, 2002).

Com o tempo, devido o avanço da tecnologia da informação e o crescente uso dos computadores, diversos sistemas especialistas foram desenvolvidos por matemáticos e engenheiros com a finalidade de facilitar as análises do mercado acionário. A tarefa de prever resultados futuros, especialmente com séries temporais econômicas, é imprescindível para análises complexas do comportamento desse mercado e por fim nas tomadas de decisões pelos investidores, buscando sempre maximizar seus lucros nas operações.(LANARI, 2000).

Todavia, ainda é um desafio para a ciência estatística e a computação desenvolver métodos e modelos que sejam plenamente eficazes para previsão do mercado financeiro, porém o uso das RNAs têm retornado valores surpreendentes, quase sempre melhores que os métodos estatísticos convencionais, por poder trabalhar com séries não lineares com muito mais facilidades e isso é comprovado pelos trabalhos aqui estudados e referenciados de (CARTACHO, 2001) e (FREITAS, 2001).

A função básica do mercado de capitais é a de promover e fomentar o crescimento da riqueza nacional por meio da participação da poupança de cada agente econômico, quais sejam investidores especialistas e mesmo o pequeno investidor dotado de pouco ou nenhum

conhecimento do mercado. Quanto mais simplificados forem os processos para a operação no mercado, maior serão os interesses das pessoas em participar e maior será o fluxo de capitais envolvidos e por consequência maior o crescimento da riqueza nacional(MENDES, 2000).

Para representar o capital das empresas, essas o dividem em ações que são títulos de créditos, denominados também por ativos financeiros e são operados nesse mercado.

2.1.1 Conceito de Ações

Uma ação consiste de um título de crédito e que, segundo a Convenção de Genebra sobre Títulos de Créditos, tem seu valor reconhecido em todos os países signatários da Convenção.

Em relação a sua causa, as ações são títulos representativos da menor fração do capital social de uma empresa (sociedade anônima, sociedade por ações ou companhia, conforme dispõe a Lei 6404/76 - Lei das Sociedades Anônimas). O detentor da ação é chamado de acionista e ele não é um credor da companhia, mas um sócio ou co-proprietário com direito na participação de seus resultados e mesmo participação em assembleias da empresa de acordo com o tipo da ação que possui.

Segundo disposto na Lei 6404/76, as ações podem ser emitidas sob a forma física de cautelas ou certificados, que comprovam a existência e a posse de certa quantidade especificada de ações, ou do tipo escritural, que dispensa sua emissão física, mantendo o controle das ações em contas de depósitos em nome de seus titulares em uma instituição depositária.

Uma ação não tem prazo de resgate, sendo convertida em dinheiro mediante negociação no mercado. O investidor pode alterar sua participação acionária, desfazendo-se de títulos possuídos ou mesmo vendendo as ações de uma empresa e adquirindo de outras.

As sociedades anônimas emitentes de ações são classificadas em: abertas ou fechadas. Uma companhia é aberta quando tem suas ações distribuídas entre um número mínimo de acionistas, podendo ser negociadas em Bolsas de Valores, para isso essas sociedades devem ser registradas na CVM (Comissão de Valores Mobiliários) como empresa de capital aberto e fornecerem ao mercado, de forma periódica, uma série de informações de

caráter econômico, social e financeiro, tratam-se das notas explicativas ao mercado e aos acionistas, oriundas de processos contábeis(HULL, 1996).

As companhias de capital fechado são tipicamente, em sua maioria, empresas familiares, com circulação de suas ações restrita a um grupo restrito de investidores e não podendo operar em Bolsa de Valores, por faltar o requisito legal constante na Lei das Sociedades Anônimasl(MENDES, 2000).

2.1.2 Tipos de Ações e sua classificação segundo a Lei das S.A.

Segundo a legislação especial conhecida como Lei das Sociedades Anônimas - Lei 6404/76 - as ações são classificadas em relação a participação nas decisões da empresa, em dois tipos: ordinárias e preferenciais.

Os acionistas que são detentores de ações ordinárias são o que comandam a assembleia de acionistas de uma empresa, conferindo ao seu titular o direito de voto. Os acionistas ordinários podem, segundo a legislação das Sociedades Anônimas:

- eleger e destituir os membros da diretoria e do Conselho Fiscal da companhia;
- reformar o estatuto social;
- decidir sobre o destinação dada aos lucros da empresa;
- autorizar emissões de debêntures (espécie de título de crédito emitido pela empresa) e aumentos de capital social;
- votar contas patrimoniais e outras ações permitidas pelo Estatuto da empresa.

Cada ação ordinária equivale a um voto, dessa forma o poder de decisão concentram-se nos investidores que detenham a maior quantidade de ações. As sociedades que tenham suas ações ordinárias concentradas nas mãos de um número reduzido de investidores têm seu controle bastante identificado com o perfil desses investidores, dificultando as livres negociações de mercado e os movimentos dos preços. Todavia, uma maior pulverização das ações, de outro modo, permite que os investidores com menor participação acionária assumam o controle da empresa, incentivando uma maior participação no capital votante(RESENDE, 2016).

À medida que o mercado de capitais se desenvolve, aumenta-se a quantidade de aci-

onistas, caracterizando uma maior distribuição das ações entre investidores e induzindo uma menor concentração no capital votante que exercerá o controle acionário de uma empresa.

As ações preferenciais são aquelas que atribuem ao seu detentor os dividendos do resultado da atividade negocial, não lhe dando nenhum direito de participação nas decisões.

2.1.3 Ações quanto a circulação

Em relação à forma de circulação, as ações podem ser nominativas, nominativas endossáveis e escriturais.

No Brasil, com a alteração da Lei das Sociedades por Ações (Lei 6404/76), a partir de 1990 no Governo de Fernando Collor de Mello, todas as ações devem ser obrigatoriamente nominativas. Anterior a isso, podiam ser nominativas, endossáveis ou ao portador, o que facilitava em muito atividades criminosas como a Lavagem de Dinheiro.

As ações nominativas trazem o nome do investidor registrado em livro próprio, denominado Livro de registro das ações nominativas. Sua transferência dá-se pela averbação do nome do novo titular neste livro, conferindo todos os direitos de acionista, bem como garante um histórico da titularidade para fins de auditorias.

Nas ações nominativas endossáveis registram-se somente o nome do primeiro acionista, sendo as transferências de titularidades processadas mediante endosso na própria cautela, processo semelhante ao utilizado pelo cheque nominal quando é transferido para outra pessoa.

As ações são emitidas por meio de cautelas, as quais devem ser registradas no referido livro de registro de ações nominativas para que se caracterize, formalmente, a propriedade do título de crédito. No entanto, existem ações anônimas sem suas respectivas cautelas, são as denominadas de ações escriturais.

Essas ações têm seu controle executado por uma instituição que figura como fiel depositária das ações da companhia, a qual as mantém em uma conta de depósito nominal aos proprietários. Todas essas movimentações com as ações ocorrem mediante extratos bancários emitidos pelas instituições depositárias.

No mercado atual a maioria das ações movimentadas são escriturais devido ao alto grau e nível de automação do sistema financeiro brasileiro.

2.1.4 Os rendimentos e riscos das ações

Ação é um título de renda variável que oferece um resultado (ganho ou perda) ao investidor baseados no comportamento de seus preços de negociação no mercado de ações. Esses resultados dependem de diversos fatores ligados ao desempenho esperado da empresa emite, das condições de mercado e evolução geral da economia do país e mundial(HULL, 1996).

Essencialmente, uma ação oferece duas formas de rendimentos aos seus investidores: os dividendos e a valorização de seu preço de mercado. Um benefício que o título também pode proporcionar aos seus titulares é o direito de subscrição em casos de aumento de capital por integralização.

Os dividendos representam distribuição de lucros auferidos pela empresa aos seus acionistas, sendo normalmente pagos em dinheiro. A decisão de uma política de dividendos envolve basicamente uma decisão sobre o lucro líquido da empresa: retê-lo, visando ao reinvestimento em sua própria atividade; ou distribuí-lo, sob a forma de dividendos em dinheiro aos acionistas.

O índice de payout (recompensa) mede a porcentagem do lucro líquido a ser distribuída aos acionistas de uma empresa e deve estar previsto formalmente no Estatuto da Empresa. Quanto maior se apresentar esse indicador de dividendos, menor será a proporção de recursos próprios, esses provenientes de decisões de retenção de lucros, utilizada pela empresa em sua atividade. Entretanto, vale lembrar que os baixos índices de payout (ou nulos) atribuem normalmente uma elevada opção para a retenção de lucros como fonte de financiamento para uma empresa, ainda pode comprometer as relações com os investidores pois esse é também um indicador visto pelo investidor para a eventual compra de ações da empresa.

Todos os acionistas (ordinários e preferenciais) têm o direito de receber dividendos em cada exercício, tornando-se uma prática obrigatória para as sociedades por ações.

A atual legislação brasileira sobre distribuição de dividendos encontra-se prevista basi-

camente na lei das sociedades por ações (Lei n.º 6.404/76). Importante frisar que os lucros são distribuídos de acordo com o previsto nos estatutos sociais e tendo como percentual mínimo obrigatório 25% sobre o lucro líquido ajustado do período em exercício.

No caso do estatuto da empresa ser omissivo com relação a esse dividendo, a legislação prevê a distribuição de 50% desse resultado ajustado para todos os acionistas.

2.1.5 Os tipos de mercado

O mercado financeiro pode ser classificado segundo diversos critérios. Considerando a natureza dos títulos negociados, tem-se o mercado de ações, o mercado de renda fixa, o mercado de créditos e outros.

Considerando-se a maturidade dos títulos são identificados principalmente o mercado monetário o qual opera em curto prazo, e o mercado de capitais com as operações de crédito de médio e longo prazos.

Uma outra classificação do mercado financeiro envolve o momento da negociação do título no mercado. O lançamento de um novo ativo financeiro ocorre no mercado primário; essencialmente, todo ativo financeiro é colocado no mercado por meio do mercado primário quando negociado a primeira vez. Dessa forma as negociações seguintes, envolvendo compras e vendas de títulos já lançados entre investidores, são realizadas no mercado secundário.

No mercado primário, ocorre a canalização direta dos recursos monetários superavitários, disponíveis aos poupadores, para o financiamento das empresas, por meio da colocação (venda) inicial das ações emitidas. É nesse setor do mercado que as empresas buscam, mais efetivamente, os recursos próprios necessários para a consecução de seu crescimento, promovendo, a partir do lançamento de ações, a implementação de projetos de investimentos e o consequente incremento da riqueza nacional. No mercado secundário são estabelecidas as renegociações entre os agentes econômicos das ações adquiridas no mercado primário. Os valores monetários das negociações realizadas nesse mercado não são transferidos para o financiamento das empresas, sendo identificados como simples transferências entre os investidores. A função essencial do mercado secundário é dar liquidez ao mercado primário, viabilizando o lançamento de ativos financeiros(STEIN E.,

1991) .

Na verdade, a existência de um mercado secundário mais dinâmico somente é possível mediante certas condições de liquidez para as ações emitidas, as quais são verificadas fundamentalmente por meio das transações nas Bolsas de Valores. Dessa forma, é nas Bolsas que uma ação, representativa de determinado empreendimento, tem seu desempenho avaliado pelos investidores, assumindo maior demanda e, conseqüentemente, maior liquidez, em função do sucesso econômico alcançado pela empresa emitente.

2.1.6 O financiamento da empresa

O financiamento de uma empresa mediante recursos próprios pode se dar pela retenção de seus resultados líquidos, o chamado autofinanciamento e também pelo lançamento de ações ao público no mercado. A colocação de ações no mercado pode ocorrer tanto pela emissão de novas ações pela empresa ou pela oferta pública de ações detidas por um determinado investidor ou grupo de investidores.

O lançamento público de novas ações, conhecido como *underwriting*, é uma operação na qual a sociedade emitente coloca seus valores mobiliários no mercado primário e para isso a faz mediante a intermediação de uma instituição financeira.

Nesse processo, a sociedade atrai novos sócios pela subscrição e integralização de novas ações (capital) e cujos recursos financeiros são direcionados para financiar seu crescimento.

Esse lançamento de ações no mercado pode ocorrer, também, pela oferta de lote de ações possuídas por um acionista (ou grupo de acionistas). Essa modalidade de oferta pública constitui-se em uma distribuição secundária de ações denominada por *block trade*. Vale ressaltar que essa operação não resulta na entrada de novos recursos à empresa, restringindo-se os resultados financeiros da operação apenas ao âmbito dos investidores. (HULL, 1998)

Essa transação secundária envolvendo um expressivo lote de ações é adotada via de regra em processos para a transferência de controle acionário perante o mercado, nesses casos um grupo controlador de uma sociedade coloca à venda as ações detidas em carteira e fora da negociação normal do mercado acionário. Ocorre bastante em operação de fusão ou de incorporação de uma empresa por outra.

2.1.7 A Bolsa de Valores

As Bolsas de Valores constituem-se, legalmente, em associações civis sem fins lucrativos, cujo objetivo básico é o de manter uma estrutura física em condições técnicas adequadas para a realização de operações de compra e venda de títulos e valores mobiliários entre os seus membros e associados.

Compete ao mercado secundário somente registrar a transferência de propriedade dos títulos e valores mobiliários, não determinando, portanto, variações diretas sobre os fluxos de recursos das sociedades. Dessa forma uma empresa somente obtém novos recursos financeiros por meio de subscrição de capital no mercado primário, dessa feita não se beneficiando diretamente das negociações do mercado secundário.

O funcionamento do mercado secundário ocorre principalmente nas Bolsas de Valores, as quais viabilizam aos investidores a oportunidade de realizarem novos negócios com títulos anteriormente emitidos.

É preocupação das Bolsas de Valores, ainda, a preservação dos valores éticos nas negociações realizadas em seu âmbito, e a divulgação rápida e eficiente dos resultados de todas as transações realizadas.

Outros objetivos das Bolsas de Valores são:

- Promover e garantir uma segura e eficiente liquidação das negociações realizadas em seu ambiente operacional;
- Desenvolver e manter um sistema seguro e regular de registro e liquidação das operações realizadas;
- Fiscalizar o cumprimento das diversas normas e disposições legais que disciplinam as operações em bolsa entre seus membros e as sociedades emissoras dos títulos.

As Bolsas de Valores são fiscalizadas e supervisionadas pela CVM – Comissão de Valores Mobiliários, atuando como uma entidade fiscalizadora do mercado de ações. São membros das Bolsas de Valores as Sociedades Corretoras que detenham título patrimonial.

2.1.8 O que é um Derivativo de Crédito

Por definição os derivativos são instrumentos financeiros que utilizam os mercados futuros e de opções principalmente com a finalidade de gerenciar riscos. Seu valor é dependente ou deriva de outros ativos aos quais eles se referem que são os ativos subja-

centes. Assim, os derivativos são contratos que refletem o comportamento futuro da bolsa de valores, do valor da moeda e da taxa de juros.

Ainda pode-se entender que derivativos são contratos que derivam a maior parte de seu valor de um ativo subjacente, taxa de referência ou índice. O ativo subjacente pode ser físico, por exemplo: Café, Petróleo, Ouro ou financeiro como Ações, Taxas de Juros, Títulos Públicos, negociado no mercado à vista ou não. Os derivativos também são criados a partir de outros derivativos.

Os contratos de derivativos podem ser classificados em: contratos futuros, opções de compra e venda, contratos a termo, operações de swaps, entre outros.

Os derivativos, em geral, são negociados sob a forma de contratos padronizados, isto é, previamente especificados (quantidade, qualidade, prazo de liquidação e forma de cotação do ativo-objeto sobre os quais se efetuam as negociações), em mercados organizados, com o fim de proporcionar, aos agentes econômicos, oportunidades para a realização de operações que viabilizem a transferência de risco das flutuações de preços de ativos e de variáveis macroeconômicas.

A negociação dos derivativos podem ser realizadas tanto no mercado primário como no secundário e também eles podem ser utilizados para quatro finalidades básicas: proteção, alavancagem, arbitragem e mesmo especulação.

Entre os tipos de derivativos destacam-se as opções, que fornecem um direito de comprar ou vender um ativo. Se a opção for de compra (call), o detentor possui o direito de comprar, se a opção for de venda (put), possui o direito de vender. O comprador de uma opção de compra, ou titular dessa opção, tem o direito de comprar certa quantidade de ações, a um preço pré-fixado até uma data determinada no contrato. O vendedor de uma opção de compra (o lançador de opção) fica com a obrigação de vendê-la. Em resumo o titular tem sempre os direitos, e o lançador sempre as obrigações.

O mercado de opções negocia esses direitos de compra ou de venda, mas não as ações diretamente. Os direitos são negociados por um preço, comumente chamado de prêmio no jargão do mercado, que é o valor pago pelo titular e recebido pelo lançador. As extinção das obrigações assumidas pelo lançador podem ser de três formas: pelo seu vencimento sem que tenha havido o exercício ou pelo encerramento (reversão) da posição; pelo exercício

da opção pelo titular. Os lançadores são obrigados a depositar garantias de forma a que se dê o cumprimento das obrigações assumidas. O valor da garantia corresponde ao dobro do valor do prêmio médio do mercado ou do seu valor arbitrado pela bolsa, sendo que essas posições margeadas terão que manter o nível até o vencimento, com os ajustes sendo calculados diariamente.

Esses derivativos são negociados no Mercado Futuro de Ações, que é o mercado no qual se negociam com datas de liquidação futura, durante o horário de funcionamento da BOVESPA, em lotes padrão nela já negociado, as ações autorizadas para tal. Cada uma das partes envolvidas, em um contrato no mercado futuro de ações da BOVESPA, se relaciona com a Companhia Brasileira de Liquidação e Custódia – CBLC, e não diretamente com a outra parte do contrato. O encerramento total ou parcial de um contrato, antes do dia de seu vencimento, também é conhecido reversão da posição de um contrato futuro.

No Brasil as negociações de futuros aconteciam na Bolsa Mercantil de Futuros BM&F, criada em de 1985, com operações a partir de 1986. Em 1991 a Bolsa de Mercadorias de São Paulo (BMSP) e a BM&F se unem criando a Bolsa de Mercadorias & Futuros, mantendo a mesma sigla: BM&F.

Em resumo o mercado futuro da BOVESPA traz a vantagem da maior liquidez, da liquidação antecipada a qualquer tempo antes da data de vencimento e a possibilidade de venda a descoberto.

2.2 Preparação de Séries Temporais

2.2.1 Conceito de Séries Temporais

Uma série temporal consiste de uma sequência de dados de uma determinada variável alocada num período. Muitos dos fenômenos científicos podem ser classificados e caracterizados como séries temporais. Na ciência econômica é possível identificar inúmeras séries temporais, tais como: o PIB, a taxa de inflação, taxa de juros, taxa de câmbio e outras (RESENDE, 2016).

Uma série temporal pode ser representada como uma sequência de dados discretos, conforme: $X_t = \{x_1, x_2, \dots, x_t\}$, onde cada valor de x_t é a representação de um valor num

dado momento, de acordo com a temporalidade da série em estudo.

Devido a grande complexidade das relações econômicas, a tarefa de estudar as séries temporais em econometria vem necessitando cada vez mais de modelos matemáticos e estatísticos, é nesta lógica que surge para o estudo das séries temporais os modelos de auto-regressão, AR (AutoRegressive - AutoRegressivo), MA (Moving Average - Média Móvel), ARMA (AutoRegression Moving Average - Média Móvel AutoRegressiva) entre outros, tem sido utilizados para previsões nos mercados financeiros e de outras séries temporais(WOOLDRIDGE, 2016).

Os conceitos seguintes são muito importantes para o entendimento do comportamento das séries temporais, são eles: tendência determinística, tendência estocástica e volatilidade na variância(BUENO, 2012).

Dessa forma, uma série temporal é dita estacionária quando a média e a variância são constantes ao longo do tempo e quando o valor da covariância entre dois períodos de tempo depende apenas da distância, do intervalo ou da defasagem entre os dois períodos de tempo, e não do próprio tempo em que a covariância é calculada.

Para o caso de uma série temporal com tendência determinística, os desvios em torno da linha de tendência, que representa a média não estacionária, são puramente aleatórios e diminuem rapidamente, não contribuindo para o desenvolvimento de longo prazo da série temporal.

Para as séries temporais com tendência estocástica, os desvios em torno da linha de tendência, que representa a média não estacionária, alteram o desenvolvimento de curto e longo prazo da série temporal. É fundamental, também, definir o conceito de volatilidade na variância, ele nos mostra se a série temporal é estacionária, vale lembrar também que nesse caso a variância não se altera com o tempo, ou é não estacionária onde a variância se altera com o tempo.

A partir dos conceitos de tendência estocástica e determinística apresentadas pode-se definir os tipos de processos de regressão. O processo auto-regressivo (AR) é aquele no qual o valor da variável dependente no período atual sofre impacto do seu valor no período anterior somado a um termo aleatório e os valores da variável dependente são expressos como desvios de seu valor médio. No processo de média móvel (MA) o valor da variável

dependente no período atual é igual a uma constante somada uma média móvel dos termos de erro presentes e passados. Já no processo auto-regressivo e de médias móveis (ARMA), a variável dependente tem características tanto do processo auto-regressivo (AR) quanto do processo de média móvel (MA).

Os modelos AR, MA e ARMA não devem ser utilizados no mercado financeiro, pois muitas vezes as séries temporais financeiras apresentam o fenômeno de aglomeração de volatilidade, ou seja, períodos com grandes variações seguidos de períodos de relativa calma. Desta feita os modelos não são adequados pois essas alterações abruptas não são passíveis de previsão nos modelos.

Para modelagem dessas séries temporais com fenômenos de aglomeração é necessário desenvolver modelos ARCH (autoregressive conditional heteroscedasticity) que é um modelo de heteroscedasticidade dinâmica no qual a variância do termo de erro, em decorrência de informações passadas, depende linearmente dos erros quadrados passados e o GARCH (generalized autoregressive conditional heteroscedasticity) que obviamente devido a sua complexidade requer um estudo bem mais aprofundado, assim as redes neurais cumprem essa deficiência dos modelos pois os processos algorítmicos podem tratar com mudanças abruptas de comportamento das séries com muito mais facilidade e eficiência, (WOOLDRIDGE, 2016).

2.2.2 Séries Temporais Estacionárias

Para que se possa fazer previsões com uma série temporal é necessário que essa seja uma série de dados ESTACIONÁRIA e através de testes estatísticos de raiz unitária é possível avaliar se uma série é ou não estacionária.

Uma série temporal é estacionária quando suas características estatísticas (média, variância, autocorrelação) são constantes ao longo do tempo. É uma série que se desenvolve aleatoriamente no tempo, em torno de uma média constante, refletindo alguma forma de equilíbrio estatístico estável (i.e. as leis de probabilidade que atuam no processo não mudam com o tempo)(BARROS, 2018).

Segundo (BARROS, 2018) existem disponíveis vários testes estatísticos de raiz unitária dentre eles:

1. Augmented Dickey-Fuller (ADF)
2. Phillips-Perron (PP)
3. Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

Na maioria dos testes a hipótese nula é de que a série tenha raiz unitária, e portanto não seja estacionária, logo:

H0 : tem raiz unitária (não é estacionária)

H1 : não tem raiz unitária (é estacionária)

Esse trabalho optou pelo teste de raiz unitária Augmented Dickey-Fuller (ADF)

2.2.3 Contornando o problema

Métodos de previsão (forecasting) usam transformações matemáticas para “estacionarizar” uma série e fazer previsões nesta série mais bem comportada, para depois inverter as transformações e obter as previsões para a série original. Transformações comuns são: tomar diferenças sucessivamente, deflacionar, aplicar log (p/estabilizar variâncias), fazer ajustes sazonais (BARROS, 2018) e (BUENO, 2012).

Como as séries temporais não estacionárias podem levar a resultados espúrios nas regressões, a literatura recomenda que os dados sejam obtidos através da primeira diferença, se essa conduzir a uma série estacionária. Nesse trabalho a série de entrada foi a série em nível, entretanto, internamente ao processo de previsão o script faz a avaliação da série quanto a estacionariedade e opera a previsão com base no vetor que se obtém com a primeira diferença (BARROS, 2018).

2.2.4 Modelo Matemático

Para os experimentos desse trabalho foi escolhido um modelo autoregressivo de ordem 5, onde cada saída é determinada em função dos valores finais da cotação dos dias anteriores, conforme abaixo:

$$y_t = x_0 + x_1 y_{t-1} + x_2 y_{t-2} + x_3 y_{t-3} + x_4 y_{t-4} + x_5 y_{t-5}$$

ou

$$y_t = x_0 + \sum_{i=1}^5 x_i y_{t-i}$$

A função para os pesos é dado por:

$$W(k+1) = W(k) + \mu(-\nabla_k E) \text{ - Equação I}$$

E foi utilizado o Gradiente do Erro Quadrado Instantâneo para estimativa obtido através da função abaixo:

$$\nabla_k E = \nabla_k e^2 = \frac{\partial f}{\partial W(k)} e^2(k)$$

Se a função de Erro é $e(k) = d(k) - y(k)$ e a função de saída da rede é dada por $y(k) = W(k)^T x(k)$ temos que:

$$\nabla_k E = \frac{\partial E(k)}{\partial W(k)} \text{ que aplicando a regra da cadeia obtemos:}$$

$$\nabla_k E = \frac{\partial E(k)}{\partial e(k)} \frac{\partial e(k)}{\partial y(k)} \frac{\partial y(k)}{\partial W(k)}$$

Desenvolvendo cada um dos fatores temos:

$$\frac{\partial E(k)}{\partial e(k)} = \frac{\partial e^2(k)}{\partial e(k)} = 2e(k)$$

$$\frac{\partial e(k)}{\partial y(k)} = \frac{\partial (d(k) - y(k))}{\partial y(k)} = -1$$

$$\frac{\partial y(k)}{\partial W(k)} = \frac{\partial (W(k)^T x(k))}{\partial W(k)} = x(k)$$

Assim temos o resultado da derivada completa agora com:

$$\nabla_k E = -2e(k)x(k) \text{ - Equação II}$$

Aplicando II em I vem a equação do peso:

$$W(k+1) = W(k) + 2\mu e(k)x(k)$$

onde:

μ é o grau de aprendizagem,

$e(k)$ é o erro na posição de execução k do vetor,

$x(k)$ é a entrada

2.2.5 Análise Exploratória e Pré-Processamento dos dados

A análise exploratória dos dados implica na verificação da qualidade e do contexto nos quais esses dados foram identificados e extraídos por processos de mineração. Para esse trabalho os dados foram disponibilizados dentro do contexto do pregão da Bolsa de Valores de São Paulo, ou seja, já está partindo de uma base de dados identificada e resultado de um processo real de ocorrência, entretanto, nem sempre isso é possível e uma etapa de mineração de dados se faz necessária, conforme (SILVA, 2016b).

É importante, nesse momento, ressaltar etapas importantes sobre os dados para que se possa obter resultados satisfatórios, assim quando a base de dados é muito grande e de impossível avaliação visual sobre sua qualidade, se faz necessário o uso de ferramentas auxiliares para essa verificação e nesse caso a Estatística Descritiva se faz presente. Os elementos da Estatística Descritiva de grande utilidade dizem respeito as medidas de posição (média aritmética, mediana e moda), separatrizes (percentis e quartis), medidas de dispersão (variância e desvio padrão), distribuição de frequência e análise de correlação.

2.3 Fundamentos sobre Redes Neurais Artificiais

2.3.1 A Inteligência Artificial e as Redes Neurais Artificiais

Os estudos que buscam reproduzir o comportamento inteligente do homem já remontam vários anos e os cientistas nessa busca acabaram por nos dar a tecnologia dos computadores, foi assim com a Máquina de Turing. (Alan Mathison Turing 1912-1954). Nasce com os trabalhos de Turing a Ciência da Computação, a qual ele é tido como o principal precursor, vem também e por conseguinte a Inteligência Artificial (IA),(RUSSEL STUART J.; NORVIG, 2003).

Segundo (RUMELHART; MCCLELLAND, 1986), no capítulo 1 do clássico livro PDP: “Esses modelos (conexionistas) assumem que o processamento de informação ocorre pela interação de um grande número de elementos processadores simples chamados de unidades, cada um enviando sinais excitatórios e inibitórios para os outros”. Para os cientistas cognitivos, porém, o apelo dos modelos de redes neurais artificiais ou conexionistas não se deve exclusivamente à sua similaridade com o cérebro. Ainda no capítulo 1 do livro PDP,

(RUMELHART; MCCLELLAND, 1986) escrevem: “Apesar de que o apelo dos modelos connexionistas é definitivamente fortalecido pela sua plausibilidade fisiológica e inspiração neural, essas não são as bases primárias do seu apelo para os cientistas cognitivos, e os modelos connexionistas chamam a atenção por razões psicológicas e computacionais. Eles contêm a esperança de que se possam oferecer descrições mecanicistas computacionalmente suficientes e psicologicamente acuradas dos fenômenos da cognição humana que têm escapado a uma explicação satisfatória via formalismos computacionais convencionais”.

2.3.2 Conceito de Redes Neurais Artificiais (RNA)

Em sua obra HAYKIN (1999, p.24) afirma que rede neural é um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela assemelha-se ao cérebro em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurônios (os pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

A base de uma rede neural são os neurônios artificiais que “simulam” o funcionamento dos neurônios biológicos de um cérebro humano. Por analogia, as entradas (inputs) para os neurônios chegam através dos dendritos, esses por sua vez também podem agir como saídas (outputs) interconectando os outros neurônios. Matematicamente, por analogia os dendritos seriam o somatório. Os axônios, por outro lado, são encontrados somente nas células de saída, cuja função sináptica é quando ativos transmitem um sinal elétrico. São responsáveis também por conectar os demais neurônios através de seus dendritos.

A característica mais relevante do uso das Redes Neurais é a capacidade de aprendizado por meio de exemplos, no qual a rede consiste de um processo iterativo de ajuste de parâmetros e está relacionado com a melhoria do desempenho da rede, sendo que o erro deve diminuir na medida em que o aprendizado prossiga. A Rede Neural possui duas fases de processamento, a de aprendizado e a de utilização ou aplicação e ambas as fases são distintas e ocorrem em momentos diferentes.

2.3.3 Estrutura das Redes Neurais Artificiais

As redes neurais artificiais são construções algorítmicas projetadas para uma determinada finalidade. É praticamente impossível criar um algoritmo que aprende com base em exemplos sem que se tenha conhecimentos de modelos matemáticos que simulem o processo de aprendizado do cérebro humano. Dessa forma basicamente, uma rede neural se assemelha ao cérebro em dois pontos: o conhecimento é obtido através de etapas chamadas de aprendizagem e pesos sinápticos são usados para armazenar o conhecimento. Uma sinapse consiste da conexão existente entre neurônios, a essas conexões são atribuídos valores, que são chamados de pesos sinápticos. Isso nos permite deduzir que as redes neurais artificiais têm em sua constituição uma série de neurônios artificiais e que serão conectados entre si, formando uma rede de elementos de processamento para o aprendizado e a execução de uma finalidade específica (CHRYSSOLOURIS, 1996).

A figura 1 ilustra um neurônio artificial com os seguintes elementos: Sinais de entrada ($x_1, x_2 \dots x_n$), Pesos sinápticos ($w_1, w_2 \dots w_n$), Combinador Linear (função soma), Limiar de ativação, Potencial de ativação, Função de Ativação ($g(\cdot)$) e Sinal de saída ($y = g(u)$).

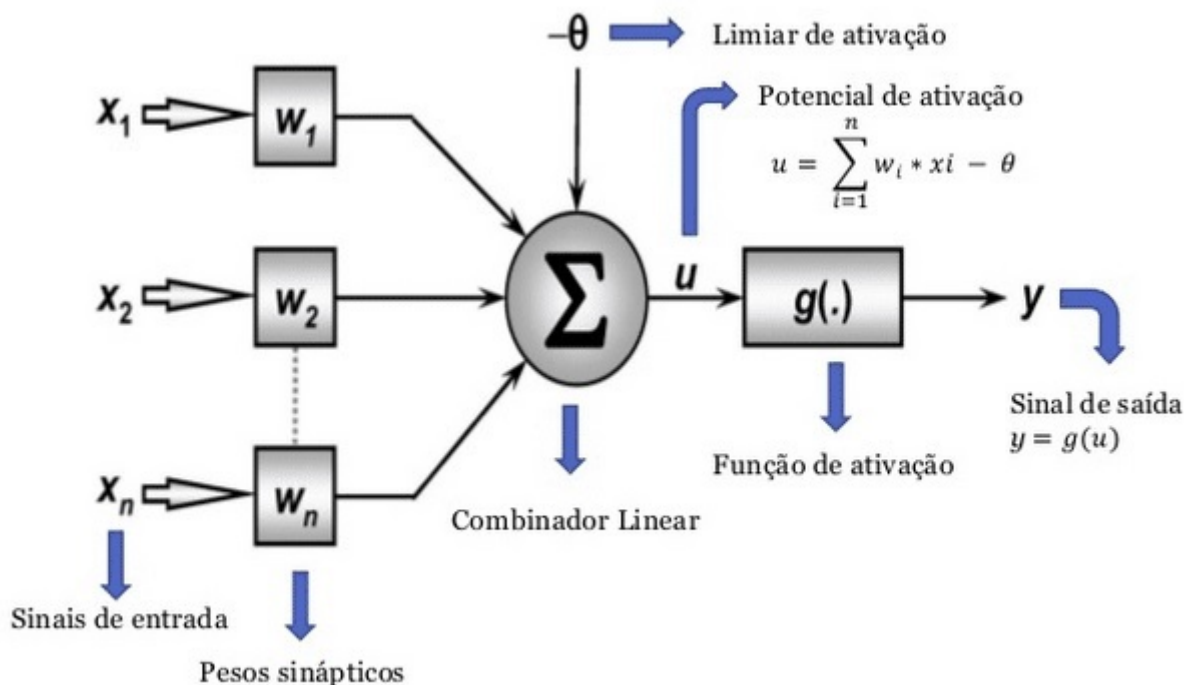


Figura 2: Neurônio Artificial - Fonte: Adaptado pelo autor de Nunes, 2016

Na figura em sua operação as entradas 1,2 e n são multiplicadas pelos respectivos pesos sinápticos e então são somadas e processadas por uma função de ativação. A função de

ativação avalia o resultado obtido de acordo com os limites definidos para depois calcular e gerar as saídas desejadas.

As funções de ativação cuja função é executar o processamento nos neurônios são: Função Ativação Linear, Função de Ativação com Limite (Threshold), Função de Ativação Sigmoid e Função de Ativação Tangente Hiperbólica dentre outras mas aqui estão as mais utilizadas, devendo ser escolhida de acordo com a funcionalidade desejada.

2.3.4 Fatores de desempenho das RNAs

As RNAs são influenciadas por três fatores a saber: (i) convergência; (ii) generalização e (iii) escalabilidade.

A convergência está relacionada ao aprendizado da RNA de identificar padrões num conjunto de dados, definido no procedimento de treinamento. A convergência mostra se o processo de treinamento é capaz de identificar os padrões não lineares numa série de dados, que será fundamental para previsões mais precisas. Esse fator merece muita atenção, pois se a rede não consegue atingir uma boa convergência, é porque não aprendeu as relações existentes e, conseqüentemente, não atingirá bons resultados na fase dos testes, devendo ter uma acurácia ruim (HERTZ, 1991).

A generalização é a etapa onde as RNAs reconhecem os padrões não existentes nos dados de treinamento, aspecto que é determinante para identificar os valores fora de padrão. Refenes (1991) afirma que a probabilidade de generalizações corretas depende do número total de redes consideradas, do conjunto de RNAs que fornecem boas soluções e do número de exemplos do treinamento. Geralmente, a redução do número de conexões da rede melhora os resultados de generalização. Porém, se deve evitar uma redução grande, que pode empobrecer e mesmo comprometer o resultado final da previsão.

A escalabilidade, que une os dois conceitos mencionados, convergência e generalização, mostra a capacidade de ajuste dos parâmetros da RNA com objetivo de convergir e generalizar de forma eficiente e satisfatória. Uma rede eficiente deve ser grande o suficiente para identificar as relações entre as variáveis, e ao mesmo tempo pequena o bastante para generalizar bem (Abelém 1994, p. 42).

Para controlar esses fatores determinantes do desempenho, é necessário escolher ade-

quadamente os parâmetros da rede que são:

- a.) Padronização dos dados
- b.) Gradiente descendente
- c.) Função de erro
- d.) Arquitetura da rede

A padronização dos dados deve ser muito bem adequada ao modelo em questão, pois dados muito ruidosos, como é o caso das séries temporais financeiras, que apresentam discrepâncias grandes entre valores ao longo do tempo, precisam ser ajustados através de técnicas como a normalização dos dados. Os dados aqui utilizados foram normalizados para o intervalo $[-1,1]$, que é fundamental para a rede ler os dados de forma adequada de acordo com a função de ativação escolhida, no caso a função de ativação sigmoidal.

O termo gradiente descendente, um algoritmo desenvolvido para identificar o mínimo local, deve também ser bem dimensionado. Azoff (1993) e Refenes (1991) destacam o uso do termo de “momentum” e da segunda derivada para reduzir o tempo de treinamento e o risco de mínimos locais, evitando erros de oscilações na rede. Outro fator importante que influencia o gradiente descendente é a taxa de aprendizado. As altas taxas de aprendizado implicam em grandes mudanças nos pesos sinápticos, reduzindo o tempo de treinamento e evitando os mínimos locais, podendo saturar o treinamento e reduzir a eficiência do processo de aprendizado; enquanto baixas taxas de aprendizado tornam o treinamento lento e aumentam a probabilidade de falsos mínimos locais.

Esse trabalho considerou a correção dos pesos sinápticos pelo Gradiente Descendente Instantâneo conforme será mostrado adiante na Metodologia.

O função de erro informa qual o grau de convergência da RNA, mostra se a rede tem capacidade de aprendizado durante o treinamento. Geralmente a função utilizada é a EMQ (Erro Quadrático Médio) ou MSE (Mean Square Error) dada pela expressão: $EMQ = \frac{1}{n} \sum_{i=1}^n (a_i - b_i)^2$ onde: a_i e b_i são respectivamente o valor real e a saída do neurônio que é o valor previsto.

Da análise dos fatores já expostos, a escolha da arquitetura da rede neural é o fator determinante de desempenho desta, assim para alcançar bons resultados de previsão é

necessário estudar as variáveis de entrada que determinam as relações do sistema em questão. Somente após a escolha ideal dessas entradas é que se define o número de conexões entre os neurônios adequadamente.

Existem alguns métodos propostos por HERTZ (1991) e REFENES (1993), como a redução de neurônios redundantes e do compartilhamento de pesos, e também o uso de um algoritmo genético, proposto por MILLER (1989), porém em nenhum deles chegou-se a bons resultados.

2.3.5 Tipos de Redes Neurais

O primeiro modelo de neurônio artificial proposto por McCulloch e Pitts (1943) era uma simplificação do que se sabia na época a respeito no neurônio biológico, onde o disparo do neurônio se fazia em função da soma dos impulsos que ele recebe até ultrapassar o limiar de excitação (threshold).

Esse comportamento era por sua vez representado no modelo artificial por um mecanismo simples, que faz a soma dos valores $x_i w_i$ (soma ponderada, onde x é o dado de entrada, w é o peso sináptico e i é interação sequencial nos vetores) recebidos pelo neurônio e decide se o neurônio deve ou não disparar (saída igual a 1 ou 0), comparando a soma obtida ao limiar ou threshold do neurônio. No modelo de McCulloch e Pitts (1943), a ativação do neurônio é obtida através da aplicação de uma função de ativação, cuja ação é ativar ou não a saída, isso dependendo do valor da soma das suas entradas.

A função de ativação é a responsável por gerar a saída Y do neurônio a partir dos vetores de peso $W = (w_1, w_2, w_3... w_n)$ e de entradas $X = (x_1, x_2, x_3... x_n)$.

A figura 2 abaixo ilustra um neurônio modelo de McCulloch e Pitts (1943).

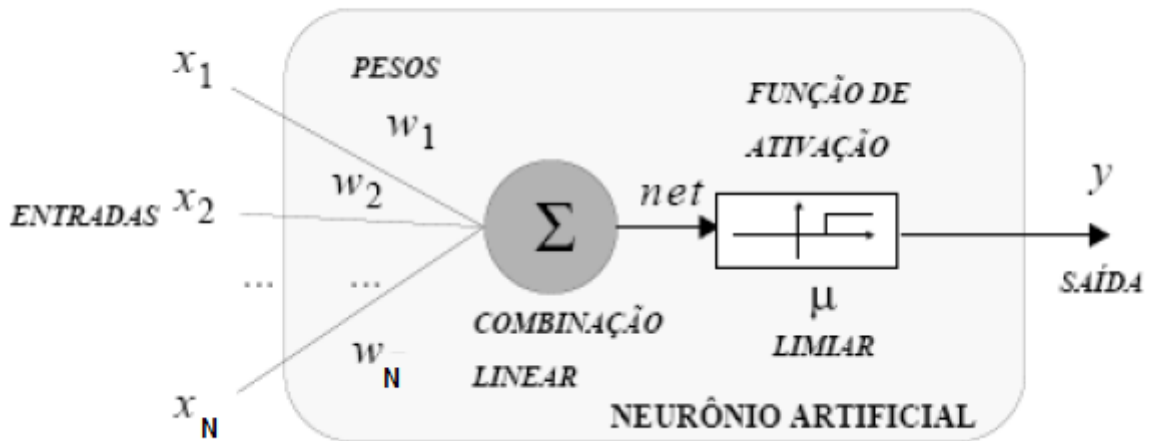


Figura 3: Neurônio de McCulloch e Pitts - 1943 - Fonte: Adaptado pelo autor de McCulloch e Pitts (1943)

2.3.5.1 Redes neurais de camada única – Perceptrons O Perceptron foi desenvolvido por Frank Roseblatt em meados de 1955 e é caracterizado por uma rede com uma única camada cujos pesos e erros podem ser treinados para se obter um vetor esperado, dado um vetor de entrada ($x(1); x(2); \dots x(n)$), onde n é o número de entradas. É o mais simples tipo de rede neural diretas (Feedforward) , conhecido como classificador linear. Isto significa que os tipos de problemas solucionados por esta rede neural devem ser linearmente separáveis.

Este tipo de rede, caracterizada por um modelo de aprendizagem supervisionada com redes de propagação para frente. Cada entrada é ponderada com peso $x_i w_i$, e a soma das entradas ponderadas é a entrada da função de transferência cuja saída é 0 ou 1.

Opcionalmente, segundo (SILVA, 2016b), um neurônio Perceptron pode ter uma entrada extra chamada bias, cujo valor é sempre 1, também chamado de valor de ativação. Na existência do bias, o algoritmo de treinamento faz o ajuste do peso b_0 , associado ao bias, com o ajuste do conjunto de pesos do vetor W da rede. A existência do bias aumenta ou diminui o sinal no campo induzido, dependendo se o valor do peso associado é positivo ou negativo, respectivamente.

O uso do bias será muito útil para tratarmos as entradas qualitativas da rede, entradas que serão obtidas com base na aplicação de lógica fuzzy sobre os resultados em

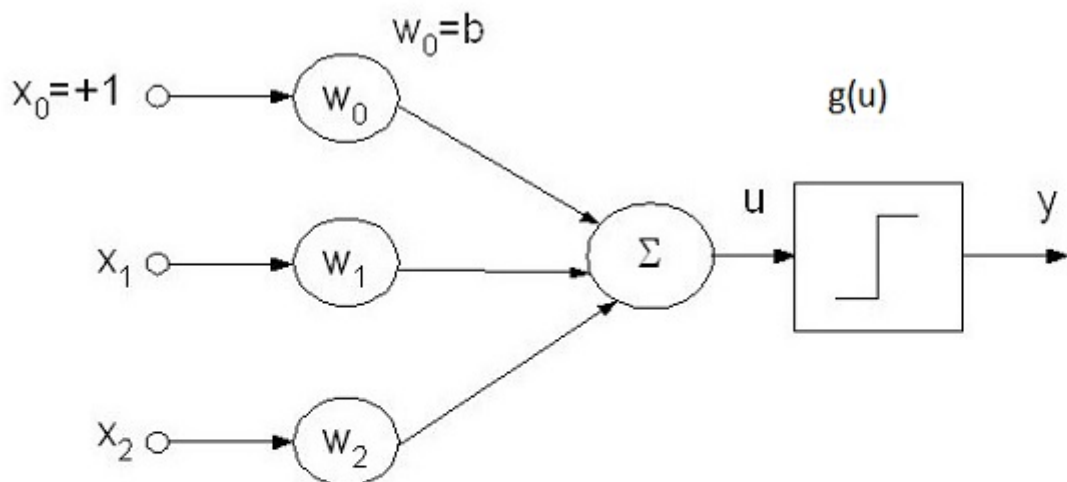
determinados ranges de datas.

A regra de aprendizado do perceptron (e) usa a diferença entre a resposta da rede (a) e o vetor de saída desejado (y), sendo y binário. Logo temos que: $e = y - a$

A cada varredura de treinamento, o perceptron tem melhor chance de calcular corretamente a saída $a = y$, dado vetor de entrada X. Os pesos iniciais são aleatórios e as regras e os aprendizados são usados repetidamente através da mudança de pesos até obtenção de um erro aceitável.

O Perceptron é um sistema de uma rede neural simples guiado pela adaptação de pesos, e onde o aprendizado é conseguido de acordo com o treino e sua capacidade de cálculo está limitada a linearidade das classes.

A figura 3 ilustra uma rede Perceptron, cuja representação é diversa das demais por se tratar de uma rede com apenas duas entradas x_1 e x_2 com o bias (opcional) representado pela entrada x_0 que assume sempre o valor +1 e pesos sinápticos w_0 , w_1 e w_2 , sendo que o peso w_0 é igual a b .



$$y = g(w_1 x_1 + w_2 x_2 + w_0), \text{ sendo } \begin{cases} g(u) = 1 & \text{se } u \geq 0 \\ g(u) = 0 & \text{se } u < 0 \end{cases}$$

Figura 4: Rede Perceptron - Fonte: Adaptado pelo autor de HAYKIN, 2001

2.3.5.2 Redes neurais lineares Neste tipo de rede também chamado de modelo Adaline (Adaptative Linear Neuron) foi desenvolvida por Widrow e Hoff em 1959.

Foi criada anos depois do aparecimento do Perceptron e é um clássico modelo de neurônio que permite saídas tanto discretas quanto contínuas e pode ser usado para tarefas de classificação e regressão.

Neste tipo de rede os neurônios têm função de transferência linear, possibilitando que a saída seja um número real e não apenas 0 ou 1.

Para estas tarefas, o Adaline apresenta uma regra de aprendizado sofisticada, a Regra Delta, que se trata de um método de regressão linear que diminuía a cada exemplo a distância entre a saída obtida e a desejada através de adaptações graduais dos pesos do Perceptron. A regressão consiste em buscar uma função que represente, de forma aproximada, o comportamento apresentado pelo fenômeno que está em estudo. A forma mais conhecida e mais usual de regressão é a linear, por exemplo, uma reta que minimiza o erro médio entre todos os valores considerados.

É inicializada com pesos aleatórios e a rede calcula os pesos de acordo com os dados de entrada e saída.

A solução é obtida pelo valor do somatório dos erros ao quadrado e a regra de treinamento é feita pelo cálculo das mudanças dos pesos de uma camada linear que façam decrescer a soma dos quadrados dos erros da camada. Calcula-se repetidamente o erro entre as entradas e saídas até que a soma do quadrado dos erros atinja um valor aceitável.

Tal regra se baseia na magnitude e no sinal do gradiente do erro para obter a direção e o valor do ajuste a ser aplicado ao vetor de pesos. O algoritmo conhecido como regra Delta deu origem anos mais tarde ao primeiro algoritmo para treinamento de redes Perceptron de múltiplas camadas, o Backpropagation.

O modelo utiliza o erro quadrático como critério de desempenho, a partir de um valor inicial para o vetor de pesos e o objetivo é tal que o vetor de pesos se aproxime gradativamente de um mínimo global, resultante de uma função Gaussiana e com superfície de formato parabólico.

2.3.5.3 Redes neurais multicamada – Multilayer Perceptrons A rede MLP (Multilayer Perceptrons) é estruturada em três tipos de camadas neurais: entrada, intermediária e saída. A camada de entrada é responsável por receber os estímulos do ambiente a ser classificado e conectá-los à camada intermediária, que têm a função de extrair a maioria das informações comportamentais da aplicação, essa por sua vez é conectada a saída que informa a que classe pertence à amostra recebida na entrada. No caso de redes com múltiplas camadas não é possível obter o erro diretamente através da diferença entre a saída desejada e a saída recorrente da rede, pois não existem saídas desejadas definidas para camadas intermediárias. O problema é calcular ou estimar o erro das camadas intermediárias(SILVA, 2016a).

Qualquer Perceptron com pelo menos uma camada escondida é um Perceptron Multicamada. Cada neurônio recebe várias entradas da camada anterior e calcula uma combinação linear dessas variáveis.

O algoritmo capaz de treinar as redes perceptron de multicamadas é o algoritmo de retropropagação ou Backpropagation, onde o princípio do seu cálculo é utilizar o gradiente descendente e estimar o erro das camadas intermediárias pela estimativa do resultado encontrado no erro da camada de saída.

O Backpropagation é um tipo de algoritmo supervisionado que utiliza pares de entrada e saída da rede para, por meio de um mecanismo de correção de erros, ajustarem os pesos da rede. O erro da saída da rede é calculado, sendo retroalimentado para as camadas intermediárias, possibilitando o ajuste desses pesos.

As redes multi-camadas têm um potencial computacional muito maior que os perceptrons de uma camada, justamente por conseguirem tratar dados não linearmente separáveis. É o número de camadas da rede que define a potencialidade de processamento da rede.

A figura 4 ilustra uma rede Perceptron Multicamadas.

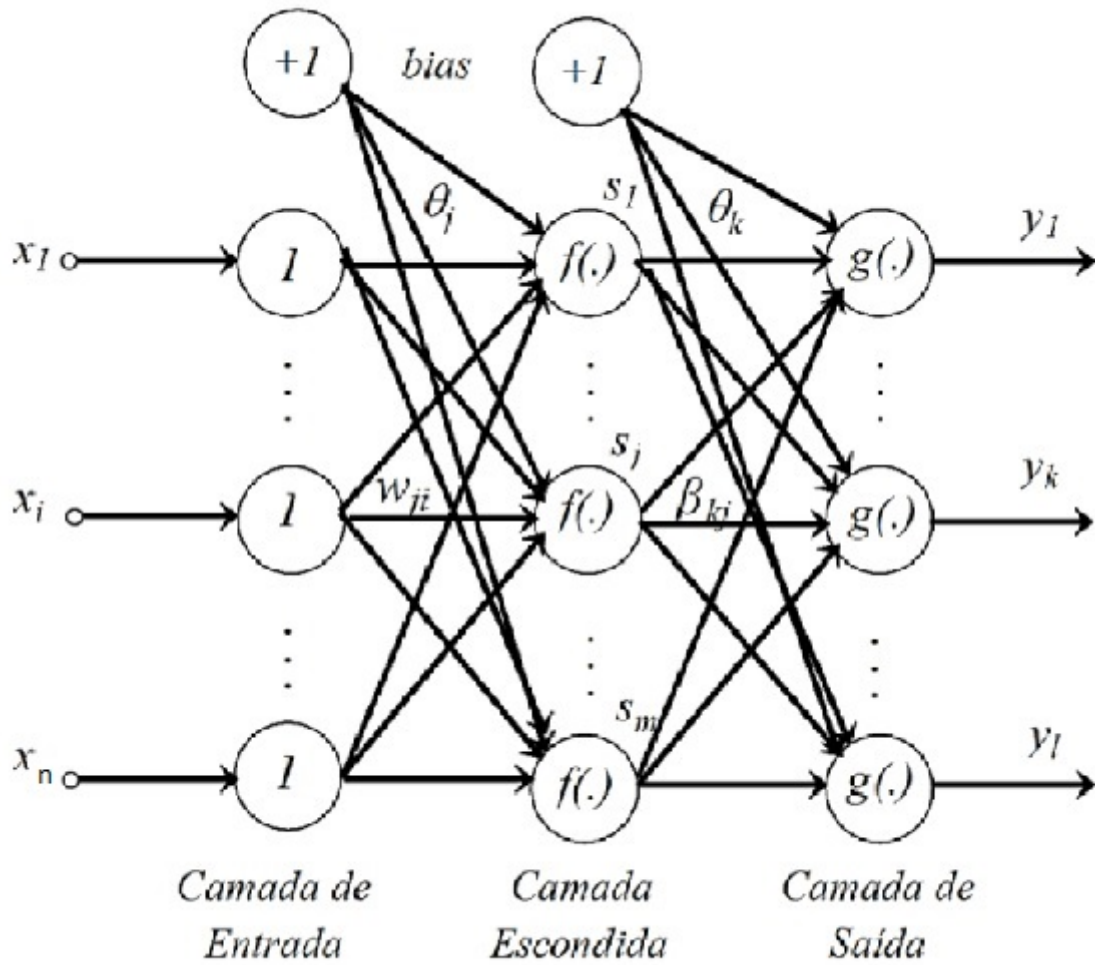


Figura 5: Rede MLP - Fonte: Adaptado pelo autor de HAYKIN, 2001

2.3.6 Tipos Algoritmos

Nos trabalhos de redes neurais existem vários algoritmos e cada um deles se adapta melhor a uma determinada atividade que se queira executar. Nesse trabalho vamos considerar o algoritmo backpropagation especificamente para o que o trabalho vai explorar.

2.3.6.1 Algoritmo Backpropagation O algoritmo Backpropagation foi criado pela generalização da regra de aprendizagem de Widrow-Hoff para redes multicamadas e não lineares.

Os vetores de entrada e os correspondentes vetores alvo são usados para treinar a rede até que se possa aproximar uma função, associando ou classificando vetores de entrada de forma adequada para os resultados específicos.

No processamento direto, uma entrada é aplicada à rede neural e seu efeito é propagado pela rede, camada a camada. Durante o processamento direto, os pesos da rede permanecem fixos. No processamento reverso, um sinal de erro calculado na saída da rede é propagado no sentido reverso, camada a camada, e ao final deste processo os pesos são ajustados de acordo com uma regra de correção de erro.

Redes com “bias” de uma camada sigmóide, e uma camada de saída linear, são capazes de aproximar qualquer função com um número finito de descontinuidades (BATTITI, 1992).

O termo Backpropagation refere-se ao modo em que o gradiente é calculado para redes multicamadas e não lineares, e tendem a dar respostas razoáveis quando se apresenta entradas que nunca foram vistas. Normalmente, uma nova entrada conduz a uma saída similar à saída correta, para vetores de entrada usados no treinamento e que sejam similares a nova entrada apresentada.

Uma das limitações do algoritmo Backpropagation é o fato dele requerer pequenas taxas de aprendizagem estável e, portanto, torna-se lento para muitos casos de aplicações práticas onde as redes são menores e com menor estabilidade. (HAGAN, 1996).

As redes multicamadas são capazes de realizar praticamente qualquer cálculo linear ou não linear, podendo aproximar qualquer função de forma bem razoável, no entanto, embora a rede treinada pode teoricamente ser capaz de executar corretamente retropropagação e suas variações, elas não podem sempre encontrar uma solução (HAGAN, 1996).

Existem outras variações do algoritmo Backpropagation que muda conforme parametrização do mesmo, e de acordo com o caminho que ele pode ser implementado. Entre eles, podemos citar: Batch Training, Batch Gradient Descent, Batch Gradient Descent with Momentun.

O algoritmo de retropropagação segue os seguintes passos:

1. Inicialização. Inicialize os pesos da rede aleatoriamente ou segundo algum critério ou método.
2. Processamento direto. Apresente um padrão à rede. Compute as ativações de todos os neurônios da rede e então calcule o erro.

3. Passo reverso. Calcule os novos pesos para cada neurônio da rede, no sentido retroativo (isto é, da saída para a entrada), camada a camada.

4. Teste de parada. Teste o critério de parada adotado. Se satisfeito, termine o algoritmo;

5. senão volte ao passo 2.

O treinamento das redes MLP com backpropagation pode demandar um tempo de treinamento longo.

2.3.6.2 Algoritmo de Random Walk O Algoritmo de Random Walk (Passeio Aleatório) é um objeto matemático que descreve um caminho, onde a série é uma sequência de passos aleatórios. Sua aplicabilidade está ligada a várias áreas do conhecimento humano, dentre eles a Economia, que é o caso desse trabalho, pois o preço flutuante de uma ação no mercado financeiro pode ser aproximado por um modelo de Random Walk.

O termo Random Walk pode se referir uma categoria particular da Cadeia de Markov, sendo essa um caso particular de processo estocástico com estados discretos (o parâmetro, em geral o tempo, pode ser discreto ou contínuo) com a propriedade de que a distribuição de probabilidade do próximo estado depende apenas do estado atual e não na sequência de eventos que precederam. A definição dessa propriedade, também chamada de memória markoviana, é que os estados anteriores são irrelevantes para a predição dos estados seguintes, desde que o estado atual seja conhecido.

Na definição o Random Walk mais simples é um caminho construído de acordo as seguintes regras:

- a.) Existe um ponto de partida ou ponto origem.
- b.) A distância euclidiana de um ponto no caminho até o próximo é constante.
- c.) A direção de um ponto no caminho para o próximo é escolhida aleatoriamente, e nenhuma direção é mais provável do que outra.

Esse trabalho considerou o Random Walk Gaussiano, onde um passo de tamanho que varia de acordo com uma distribuição normal é usado como um modelo para séries de dados temporais do mundo real, tais como os mercados financeiros. A fórmula de

Black–Scholes para a modelagem de opção de preços, por exemplo, usa um Random Walk Gaussiano como a suposição subjacente. Aqui, o tamanho do passo é o inverso da distribuição cumulativa normal.

Ainda, é importante ressaltar o mérito da fórmula de Black-Scholes que se baseia no conceito da equação de difusão que é uma equação em derivadas parciais que descreve flutuações de densidade em um material que se difunde. É também usada para descrever processos exibindo um comportamento de difusão, aqui considerado o movimento das ações no mercado financeiro. A equação de difusão é uma generalização da equação do calor base dos trabalhos de Fourier (Jean-Baptiste Joseph Fourier (1768-1830)) e que tem grande aplicação na estatística.

2.3.7 Modelos Utilizados

Os modelos utilizados nesse trabalho foram: Rede Neural MLP, com os algoritmos Backpropagation com função de ativação Sigmoid e Random Walk aplicado a série temporal, lembrando que esse não se trata de uma rede neural.

3 METODOLOGIA

Nessa parte do trabalho será apresentada a caracterização da pesquisa e dos devidos procedimentos metodológicos. Em linhas gerais, segundo PARRA Domingos Filho; SANTOS (2002), pode-se classificar a presente pesquisa como pesquisa aplicada, uma vez que visa atingir resultados imediatos. Contudo, conforme o autor SANTOS (2002), pode-se caracterizar a metodologia empregada nos trabalhos científicos, focando em três critérios principais: objetivos, procedimentos de coleta de dados e fonte das informações.

O ponto de partida para o presente trabalho deu-se pela pesquisa bibliográfica que resgatou modelos matemáticos que são utilizados na análise técnica do mercado de capitais brasileiro, focados no que é ensinado na disciplina de Econometria, bem como os estudos aplicados em redes neurais artificiais MLP que é o objetivo comparativo desse trabalho.

O presente trabalho contou com os seguintes procedimentos metodológicos para sua consecução, conforme os passos ilustrados na figura 6. Cada um dos procedimentos será

apresentado em detalhes no decorrer deste capítulo.

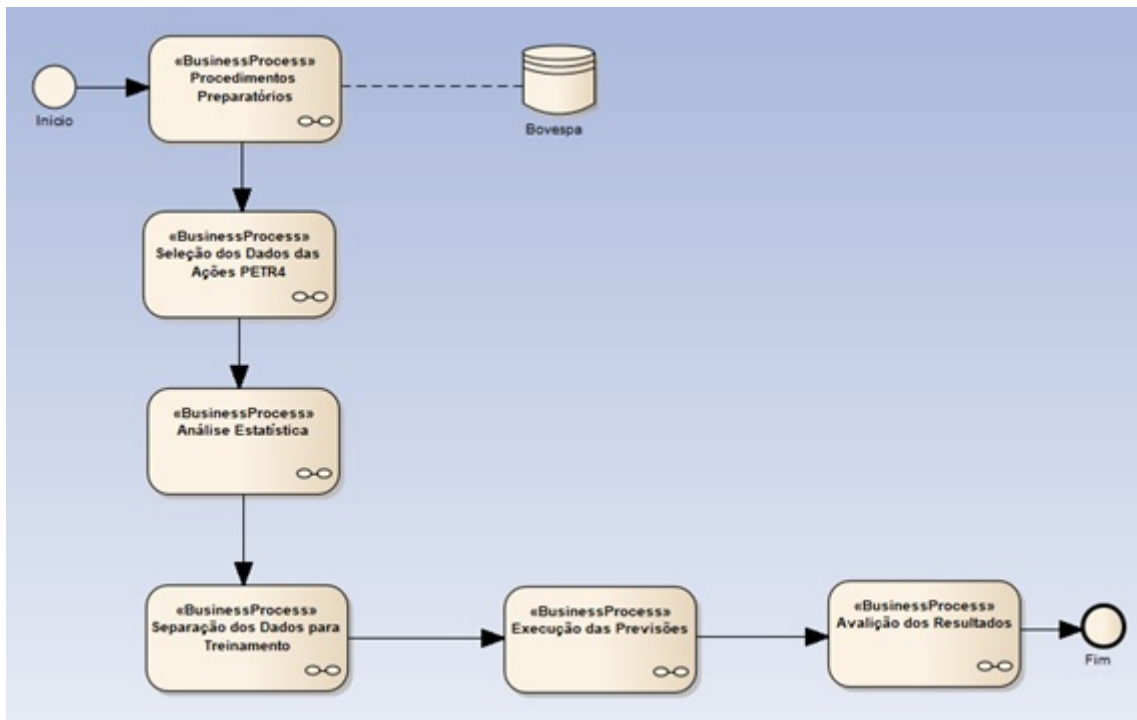


Figura 6: Fluxo de execução da metodologia experimental - Fonte: Autor

3.1 Procedimentos Preparatórios

A preparação dos dados é um procedimento importante para a realização dos experimentos a serem executados. Esta etapa consiste na coleta dos dados que será feita a partir dos dados disponíveis de forma livre no website da Bolsa de Valores de São Paulo, BOVESPA.¹ Os dados coletados em forma de arquivos com informações sobre as ações negociadas no período de 1995 a 2017, num total de 5.007.516 registros das diversas empresas que compõe o pregão diário da operação de mercado BOVESPA.

Os dados coletados foram armazenados em tabelas relacionais em um banco de dados SQL Server de onde foram amostrados para a realização dos experimentos desta pesquisa, definindo assim períodos específicos de análise para comparação de desempenho de predição da rede neural e demais modelos preditivos. Justifica-se ainda para uso de tabelas relacionais o número de informações que são disponíveis aos dados brutos, trazendo va-

¹http://www.bmfbovespa.com.br/pt_br/servicos/market-data/historico/mercado-a-vista/series-historicas/

riáveis que não são de interesse deste estudo. Portanto, os dados armazenados em um banco se tornam mais fáceis de serem manipulados.

3.2 Seleção dos dados das ações PETR4

De toda a base de dados coletada da BOVESPA um total de 4456 registros são das ações PETR4 da empresa Petrobrás, que será o objeto de estudo deste trabalho. Os dados selecionados para análise referem-se as datas das operações e preços do efetivo fechamento do pregão diário. Sendo assim, a sequência de dados disponíveis das ações PETR4 representa uma série temporal, da qual deve-se antes de realizar algum tipo de modelagem preditiva passar por uma análise estatística em que se analisa o comportamento da série quanto a presença ou não de raiz unitária.

3.3 Análise Estatística

Para que uma série temporal possa ser analisada por uma modelagem preditiva é necessário que tenha comportamento de “estacionariedade”, do qual garante confiança nos resultados de predição. A análise estatística que auxilia na detecção deste comportamento investiga a série no sentido de identificar a presença de tendências, seja essa determinística ou estocástica. Como discutido no capítulo 2 a análise de tendência é feita com o teste de raiz unitária, do qual existe na literatura diferentes métodos, que para este trabalho optou-se por utilizar o Teste Augmented Dickey-Fuller. No teste ADF estabelecem duas hipóteses:

- H_0 : tem raiz unitária (não é estacionária)
- H_1 : não tem raiz unitária (é estacionária)

É importante ressaltar que em sendo a série não estacionária, ou seja, o resultado do teste H_0 , não significa que a modelagem da série seja inviável, mas sim na necessidade de se fazer uma transformação na mesma série. Geralmente a transformação é feita a partir da primeira diferença, ou aplicando a função logaritmo. A transformação implica em se ter uma nova série a qual também é denominada como série em nível, e que pode ser submetida a função inversa da transformação que foi submetida, por exemplo, em caso a nova série seja gerada por um logaritmo, a obtenção da série original é feita a partir

da aplicação de uma função exponencial. Após a transformação na série, deve-se analisar novamente a “estacionariedade” e verificar a existência ou não da raiz unitária. Sendo a nova série estacionária essa série que será usada na previsão, conforme explicado no capítulo 2.

Adiante no item 5 que trata dos experimentos e resultados têm-se a aplicação do teste ADF na série em nível e na primeira diferença da série PETR4 a ser utilizada nesse trabalho.

3.4 Separação dos dados para treinamentos

Para o processamento dos experimentos foram selecionados dois lotes de dados da ação PETR4, sendo o primeiro lote contendo 2.223 registros do período de 02/01/2000 a 30/12/2009, que trataremos como cenário 1 e outro lote de 1973 registros do período de 04/01/2010 a 28/12/2017 que trataremos como cenário 2. É importante consignar que os períodos selecionados compreendem dois fatos importantes ocorridos no país e que tiveram efeitos significativos no mercado financeiro brasileiro, o primeiro cenário consiste na Crise Americana de 2008 e o segundo a recessão que atingiu o Brasil no começo do ano de 2015. O objetivo dos cenários é configurar situações de grande flutuação nos preços das ações, dentre elas a PETR4 e avaliar o uso de modelos preditivos na predição correta. O dado considerado na série de cada cenário, como falado antes, é o Valor do Fechamento da ação na data do pregão. Conforme já explicado no capítulo 2, estes dados serão organizados seguindo um modelo auto regressivo de ordem 5, conforme ilustrado na figura 7, ou seja: as entradas x_1 a x_5 são os valores do fechamento em dias consecutivos, por exemplo: x_1 é o valor do fechamento em 02/01/2001, x_2 é o valor do fechamento em 03/01/2001, x_3 é o valor do fechamento em 04/01/2001, x_4 é o valor do fechamento em 05/01/2001, x_5 é o valor do fechamento em 08/01/2001 e a saída esperada y é o valor do fechamento em 09/01/2001 e assim a série é estruturada para uso na rede neural, sendo o exemplo representando uma amostra do conjunto de dados.

Os dados estruturados em cada um dos cenários serão separados em dois conjuntos, treinamento e teste. A proporção escolhida para cada conjunto é de aproximadamente 70% de treinamento e 30% de teste, cuja escolha das amostras para compor cada um dos conjuntos é feita de forma aleatória. Na figura 8 estão os metadados dos conjuntos de

Estrutura da Série de Entrada						
Índice	X1	X2	X3	X4	X5	Y
1	47,09	49,65	50	49,53	49,39	50,43
2	49,65	50	49,53	49,39	50,43	50,09
3	50	49,53	49,39	50,43	50,09	50,8
4	49,53	49,39	50,43	50,09	50,8	51,3
5	49,39	50,43	50,09	50,8	51,3	51,3
6	50,43	50,09	50,8	51,3	51,3	52,43
7	50,09	50,8	51,3	51,3	52,43	53,38
8	50,8	51,3	51,3	52,43	53,38	54,1
9	51,3	51,3	52,43	53,38	54,1	55,27
10	51,3	52,43	53,38	54,1	55,27	54,6
11	52,43	53,38	54,1	55,27	54,6	55,81
12	53,38	54,1	55,27	54,6	55,81	55,1
13	54,1	55,27	54,6	55,81	55,1	54,35
14	55,27	54,6	55,81	55,1	54,35	54,4
15	54,6	55,81	55,1	54,35	54,4	54,45
16	55,81	55,1	54,35	54,4	54,45	54,51
17	55,1	54,35	54,4	54,45	54,51	53,8
18	54,35	54,4	54,45	54,51	53,8	53,8
19	54,4	54,45	54,51	53,8	53,8	53,6

Figura 7: Visão parcial da série de dados de entrada na rede neural - Fonte: Autor

dados para os dois cenários.

Distribuição das Amostras					
Cenário	Amostras	Treinamento	Previsão	% Treinamento	% Previsão
1	2223	1523	700	68,51%	31,49%
2	1973	1373	600	69,59%	30,41%

Figura 8: Distribuição quantitativa das amostras - Fonte: Autor

3.5 Execução das previsões

As previsões serão executadas em um total de 16 (dezesesseis) execuções. Para o processamento dos dados será utilizado o software R-Studio com as bibliotecas neuralnet e tseries que implementam as redes MLP e a Hipótese de Random Walk. (a citação das bibliotecas faço na bibliografia, mapeando o código) As dezesseis execuções serão feitas com

a rede neural MLP para os dois cenários considerados. Serão executados experimentos de variação de parâmetros da rede, especificamente no número de neurônios na camada intermediária, considerando configurações para redes com 5 e 10 neurônios. Adicionalmente ao experimento de neurônios na camada intermediária serão executados experimentos na parametrização de valores de taxa de aprendizado, considerando valores de 0,1 a 0,8, com incremento de 0,1. Com a finalidade de evitar eventuais vies de resultados de predição, as amostras dos conjuntos de dados serão formadas aleatoriamente. Sendo assim, os resultados serão apresentados em termos de média e desvio padrão.

Por fim, o melhor resultado obtido em cada cenário será comparado com o resultado obtido por Random Walk.

3.6 Avaliação dos resultados

Para a análise dos resultados obtidos será utilizado o indicador de Erro quadrático médio (RMSE), também conhecido como Risco quadrático, que é definido como sendo a média da diferença ao quadrado entre o valor estimado e o valor real presente na base de dados.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$$

Será considerada no quadro comparativo aquela que apresentar o menor valor para o RMSE.

4 EXPERIMENTOS E RESULTADOS

Neste capítulo serão explicados os experimentos realizados, iniciando-se, com a preparação efetiva das séries de dados que serão utilizadas para a modelagem preditiva da rede neural e random walk. Como já explicado no capítulo 3 e lembrado no capítulo 4, é imprescindível que a série em nível seja uma série temporal estacionária. Nas figuras 9 e 10 tem-se os gráficos das séries a serem utilizadas, respectivamente nos cenários 1 e 2. Em cada figura seguem com valores em nível, conforme coletada da base de dados BOVESPA. Da figura é possível verificar certa tendência no comportamento dos preços

das ações PETR4 e os valores das ações estão distantes do gráfico da média dos valores da série, disso é possível perceber o comportamento não estacionário da série, o que não é recomendado para previsão.

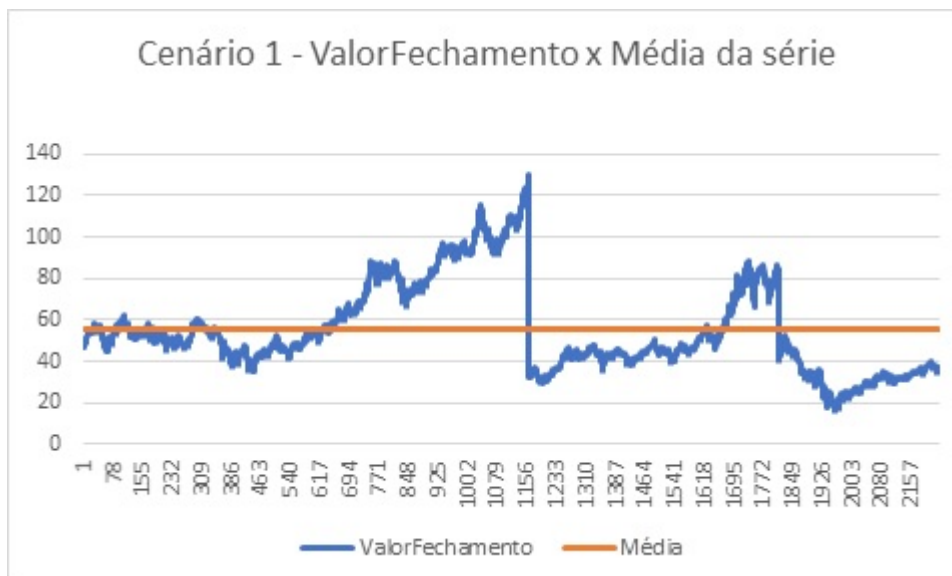


Figura 9: Cenário 1 - Série em nível - Fonte: Autor

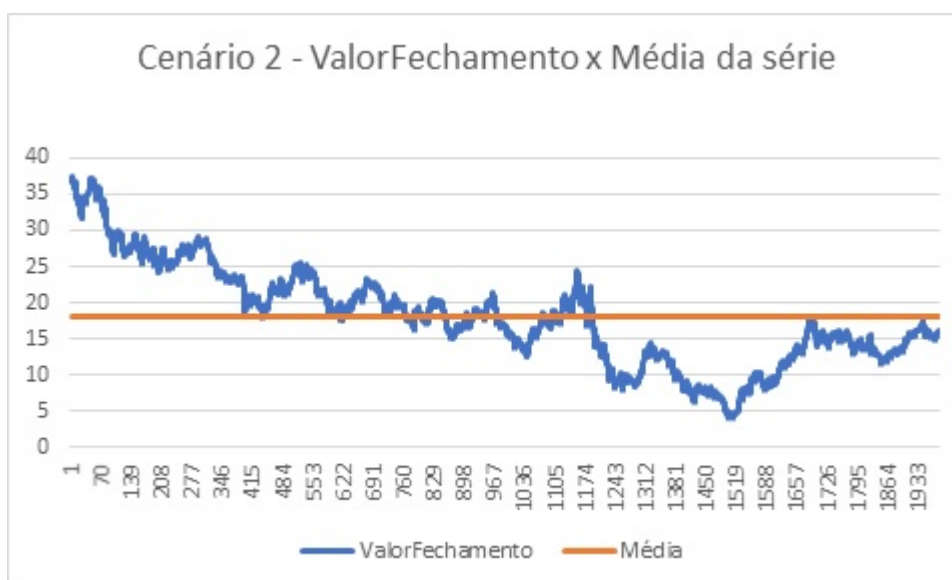


Figura 10: Cenário 2 - Série em nível - Fonte: Autor

Embora as figuras acima indiquem a situação de estacionariedade das séries de nível, a certificação das hipóteses de que as séries apresentadas são não estacionárias só podem ser verificadas com a execução do teste de raiz unitária ADF. O resultado completo deste teste, para ambas as séries, estão apresentados na Tabela 11.

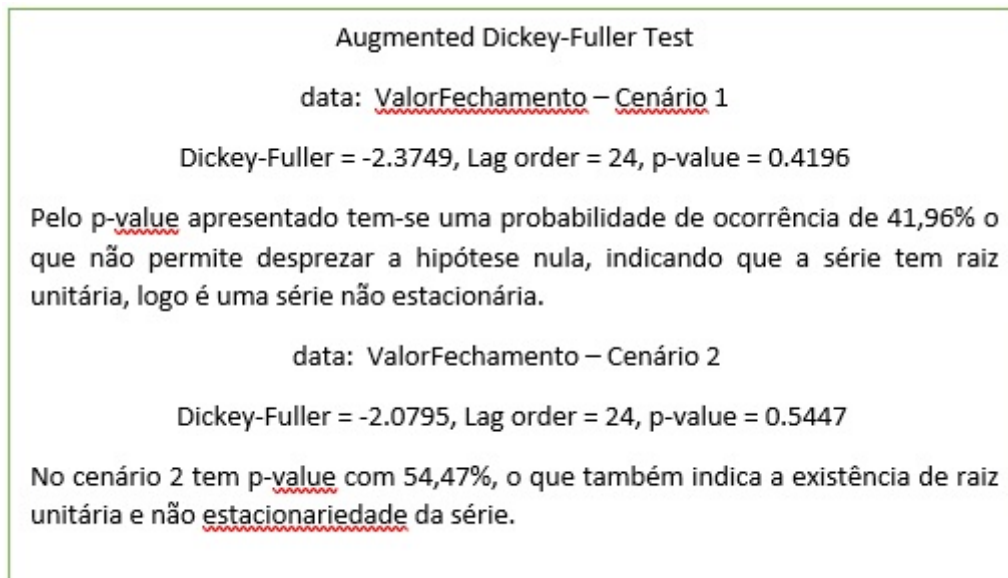


Figura 11: Resultado dos testes ADF sobre os dados da série em nível

Como se observa do teste ADF apresentado na Tabela 11, confirma-se a indicação gráfica de que as séries não são estacionárias. Com isso se faz necessário a aplicação da primeira diferença nas séries de níveis, cujos resultados estão nas figuras 12 e 13. Nas séries de diferença pode-se verificar que a série tem um comportamento muito próximo da média da série. No entanto, mesmo assim, não é condição suficiente para a afirmação da estacionariedade, o que será obtido também com o teste ADF, mostrado na tabela 14.

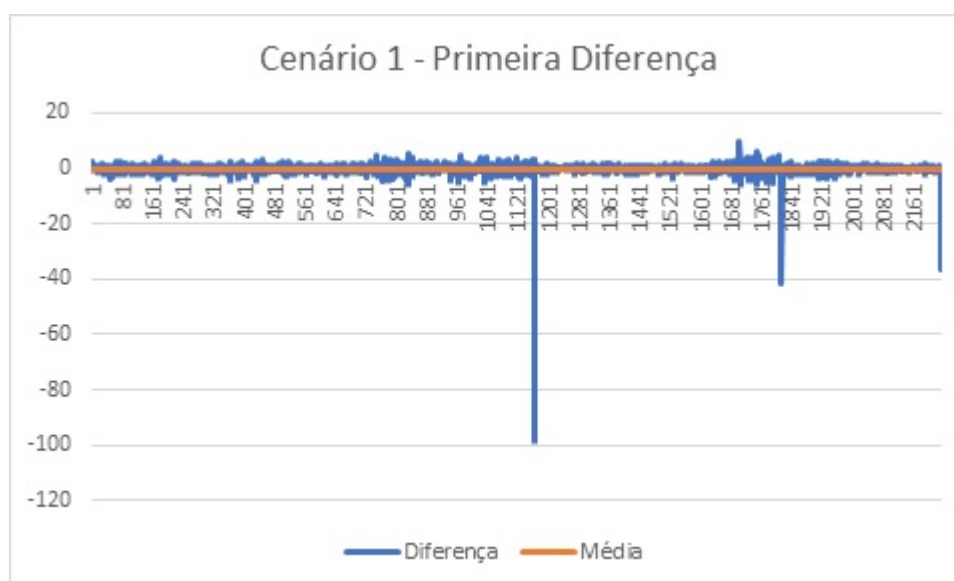


Figura 12: Cenário 1 - Série da primeira diferença - Fonte: Autor

Da análise da aplicação do teste sobre a primeira diferença obtém-se um p-value muito

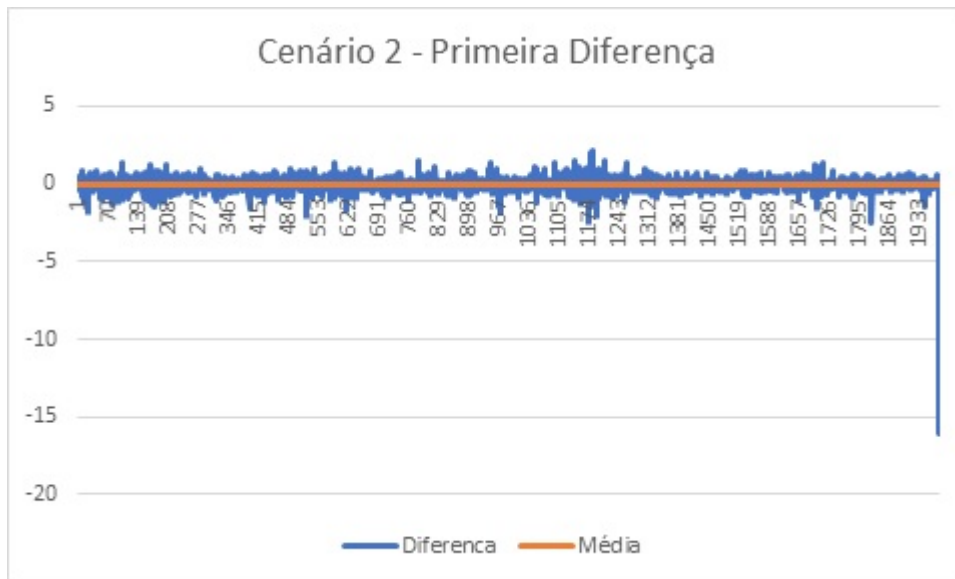


Figura 13: Cenário 2 - Série da primeira diferença - Fonte: Autor

Augmented Dickey-Fuller Test

data: Diferença – Cenário 1

Dickey-Fuller = -9.8437, Lag order = 24, p-value = 0.01

p-value smaller than printed p-value

data: Diferença – Cenário 2

Dickey-Fuller = -7.0021, Lag order = 24, p-value = 0.01

p-value smaller than printed p-value

Na saída do teste com a primeira diferença da série, o software que executou o processamento alerta para um p-value muito menor do que o exibido, recomendando ainda mais a hipótese de estacionariedade da série da primeira diferença.

Figura 14: Resultado dos testes ADF sobre os dados da série da primeira diferença

menor do que 5% ($p\text{-value} < 0,05$) o que segundo a literatura e visto no capítulo 3, permite desprezar a hipótese nula, obtendo-se assim uma nova série estacionária e própria para a previsão.

O p-value é um indicador estatístico sobre a hipótese nula num teste de hipótese, valores em torno de 5% ($p\text{-value} \leq 0,05$) recomendam que se pode rejeitar a hipótese nula. A hipótese nula sendo rejeitada indica que a série é estacionária.

Os experimentos executados são iniciados com os dados da série em nível, todavia, como já enunciado na metodologia, a série será submetida a uma análise de estacionariedade antes da efetiva entrada na rede neural e isso é feito no script de execução que processa o teste ADF e decide pelo resultado obtido no p-value se há a necessidade de fazer a primeira diferença e em seguida gerar a série para a entrada na rede neural. Importante ressaltar que os resultados da previsão usando a primeira diferença, cujos gráficos das séries de entradas, comparativo com as médias, encontram-se nas figuras 12 e 13 para os dois cenários considerados, são somados com o valor da série em nível para se obter o resultado da previsão.

Definidas as amostras e aplicados os devidos preparos necessários à série de dados, foram executados 96 experimentos, sendo 48 experimentos para uma rede neural MLP com 5 neurônios e 48 experimentos para a rede com 10 neurônios, conforme resultados exibidos na figura 15. Os experimentos foram divididos em ensaios, sendo 3 ensaios com 16 experimentos para o cenário 1 e o mesmo para o cenário 2.

Na figura tem-se a taxa de aprendizagem variando de 0,1 a 0,8 para os experimentos de 1 a 8 que correspondem a rede com 5 neurônios e os de 9 a 16 correspondem a rede com 10 neurônios.

Da análise da figura conclui-se que os experimentos assinalados em cor azul são os de menor erro para cada um dos cenários e serão utilizados para a comparação com a Hipótese de Random Walk, conforme figura 17.

Comparativo do Erro RMSE						
Experimento	Aprendizagem	Cenário 1			Cenário 2	
		Ensaio 1	Ensaio 2	Ensaio 3	Ensaio 1	Ensaio 2
1	0,1	0,058551591	0,042836669	0,042362305	0,101764429	0,083374642
2	0,2	0,042313143	0,05198354	0,044796317	0,12194843	0,128065184
3	0,3	0,063443719	0,050832343	0,034774617	0,103494442	0,049493828
4	0,4	0,040915905	0,045924498	0,058067845	0,076839356	0,088256823
5	0,5	0,077938818	0,066863646	0,05978828	0,034485209	0,1182103
6	0,6	0,05003226	0,052525148	0,03541763	0,156563689	0,088329597
7	0,7	0,047562623	0,068300261	0,109480242	0,043219656	0,042860963
8	0,8	0,024568012	0,063112577	0,059344814	0,036207358	0,0491588
9	0,1	0,037113951	0,050127702	0,062496061	0,037366174	0,244406467
10	0,2	0,031482577	0,073207188	0,071717535	0,216906606	0,060902846
11	0,3	0,050876613	0,043977126	0,058598704	0,030655995	0,108421146
12	0,4	0,050670772	0,08764733	0,066787887	0,112216624	0,115514815
13	0,5	0,041700956	0,046368323	0,063881283	0,099782092	0,034772408
14	0,6	0,036401855	0,04477888	0,045519085	0,09905774	0,067358856
15	0,7	0,048060039	0,082370994	0,066581731	0,130945922	0,167226712
16	0,8	0,071206941	0,032950228	0,048002353	0,07036587	0,028319848

Figura 15: Quadro comparativo do Erro MSE para vários ensaios por cenário

Legenda	
	Menor Erro com 5 neurônios
	Menor Erro com 10 neurônios
	Menor Erro com 5 e 10 neurônios

Figura 16: Legenda

Resumo dos Experimentos com a Rede MLP			
Experimento	Nr.Neurônios	Taxa Aprendizado	Erro RMSE
8	5	0,8	0,024568012
15	10	0,7	0,024433156

Figura 17: Experimentos selecionados para comparação com Random Walk

Nas figuras 18 e 19 são mostrados o comportamento do erro (em ordenada) em função da taxa de aprendizado (em abcissa) para a rede neural MLP, com 5 e com 10 neurônios.

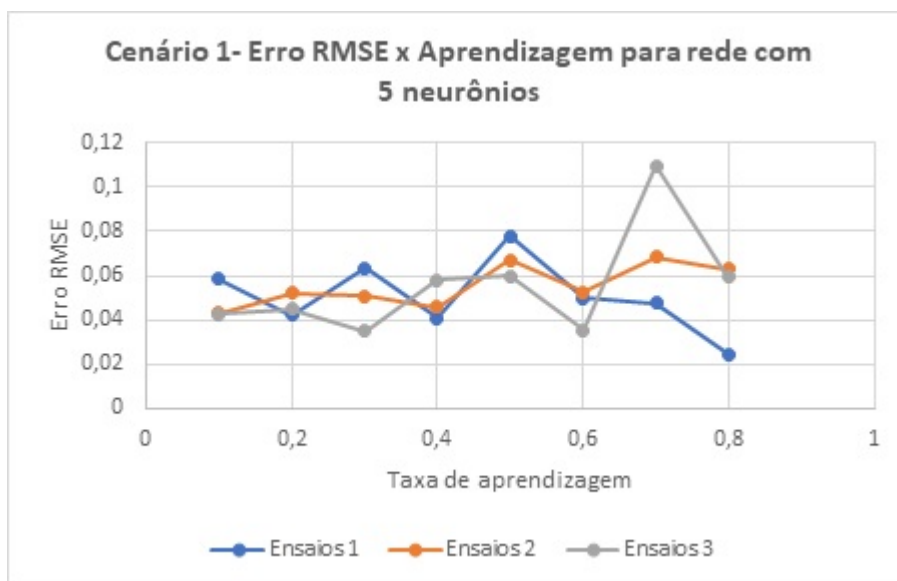


Figura 18: Erro x Taxa de aprendizagem para experimentos com 5 neurônios

Nas figuras 20 tem-se o gráfico do processamento da previsão para o experimento 8 com 5 neurônios e 0,8 como taxa de aprendizado. Do gráfico percebe-se uma aproximação entre os valores previstos e observados. Na figura 21 é mostrada a diferença entre o valor previsto e o valor observado, nele percebe-se uma aproximação preponderante do eixo de valor zero, todavia há um ponto de inflexão na série que provocou um salto no gráfico. Entretanto, o comportamento da rede foi o de acompanhar a inflexão, inclusive

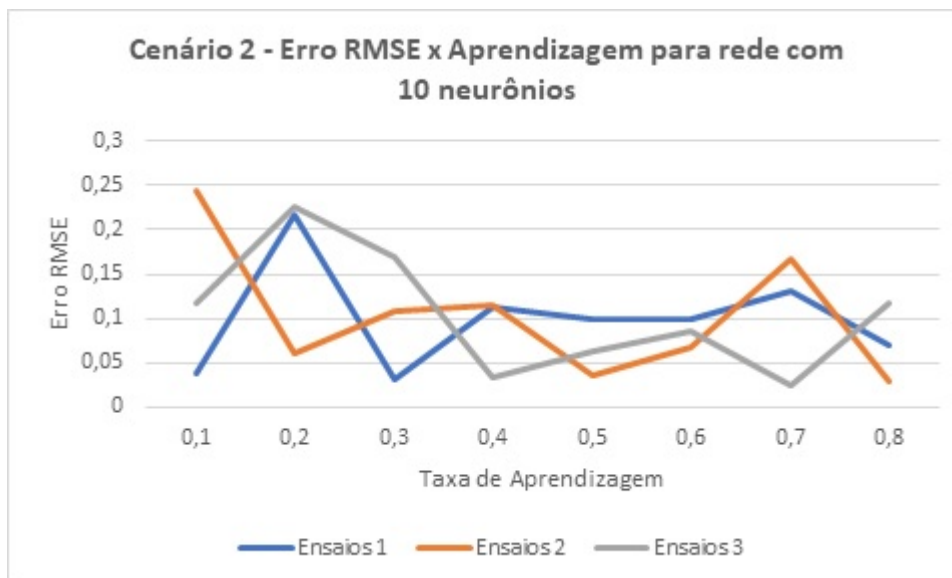


Figura 19: Erro x Taxa de aprendizagem para experimentos com 10 neurônios

com valores muito próximos. Isso mostra que a rede neural é eficiente para tratar pontos de grande variação nos valores informados. Por esse comportamento denota-se que o processo de previsão com o uso de rede neural é satisfatória e acompanha bem as séries com razoáveis flutuações nos preços.

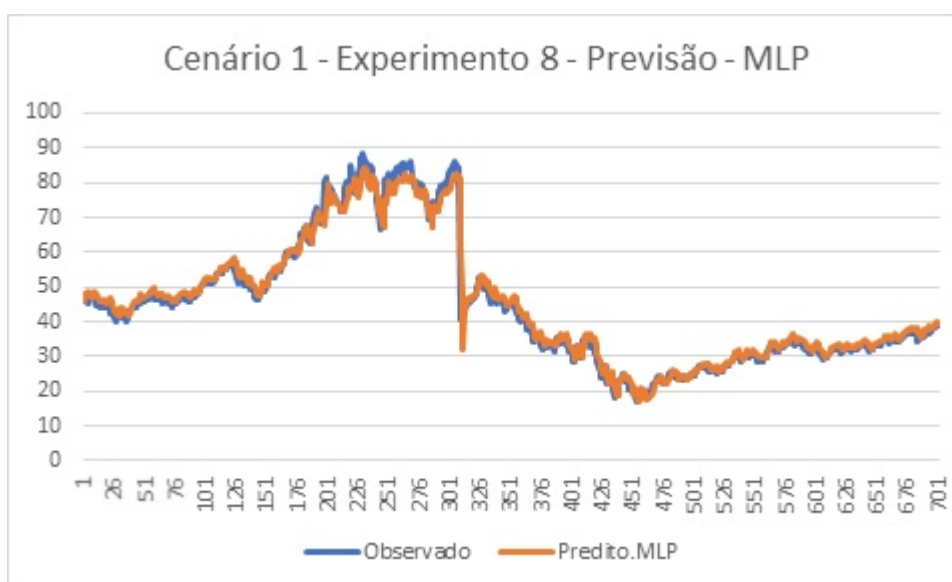


Figura 20: Cenário 1 - Experimento 8 - Previsão - MLP

Na figura 22 mostra-se os valores observados e os previstos pela rede MLP nesse experimento. A coluna Diferença MLP é o resultado da diferença entre os valores previstos e observados.

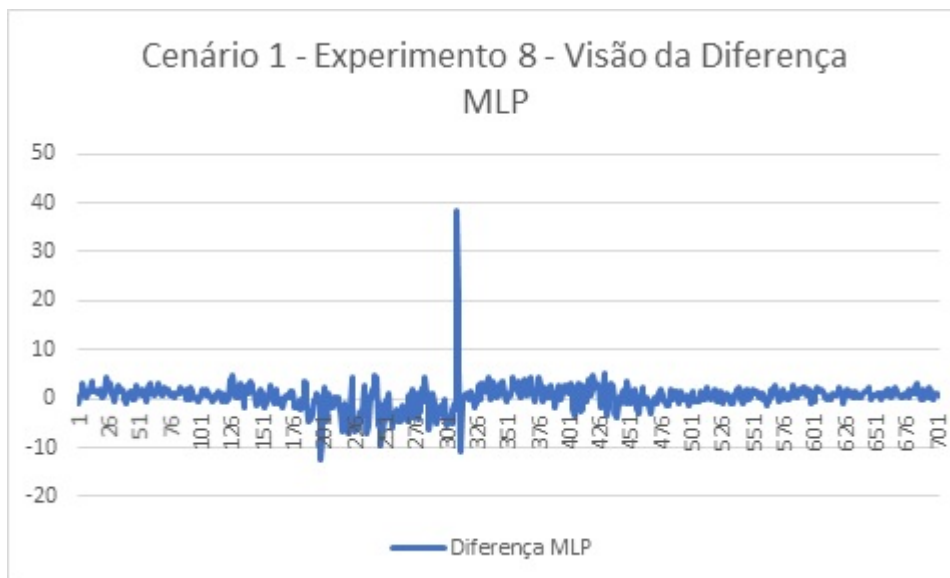


Figura 21: Experimento 8 - Diferença previsto e observado - MLP

Indice	Observado	Predito.MLP	Diferença MLP
1	46,7	45,52142155	-1,178578447
2	47,1	47,87620468	0,776204678
3	46,5	47,6136116	1,113611596
4	45,27	48,55348506	3,283485061
5	46,31	47,65863125	1,348631248
6	46,87	47,27461611	0,404616107
7	46,74	46,97472335	0,23472335
8	47,3	48,30163554	1,001635536
9	47	48,54663584	1,546635836
10	46,26	48,20146165	1,941461655
11	44,91	48,31577251	3,405772506
12	45,6	47,14895904	1,548959045
13	44,8	46,75743098	1,95743098
14	44,25	45,79978326	1,549783256
15	45,06	46,46038397	1,400383966
16	44,75	45,83183389	1,081833892
17	44,24	45,72642352	1,486423518
18	44,33	46,29482	1,964819996
19	44,3	45,68515341	1,385153409
20	45,05	45,41722636	0,367226361
21	45,43	45,9988105	0,568810495
22	45,61	46,22969596	0,619695955

Figura 22: Cenário 1 - Experimento 8 - Dados da Previsão - MLP

Nas figuras 23, 24 e 25 tem-se a mesma demonstração para o experimento 15 que será também utilizado para comparação.



Figura 23: Cenário 2 - Experimento 15 - Previsão - MLP

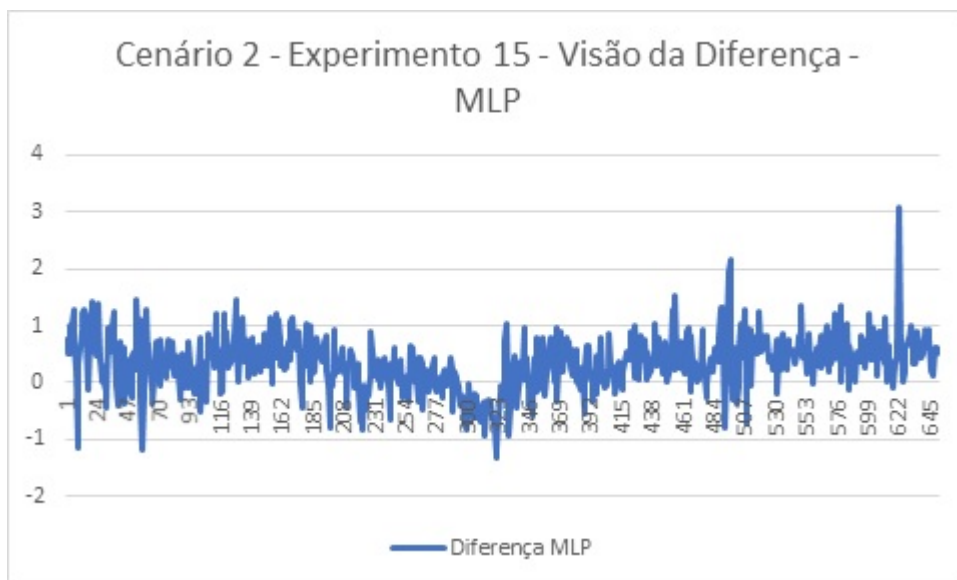


Figura 24: Cenário 2 - Experimento 15 - Diferença previsto e observado - MLP

Conhecidos os dois experimentos, 8 e 15, conforme figura 17, os quais apresentaram o menor valor para o erro RMSE, esses dados serão utilizados para comparação com a Hipótese de Random Walk nas mesmas condições de processamento, no caso Random Walk usando a mesma série de dados e no caso da MLP conforme a parametrização mostrada a seguir.

É importante ressaltar que a rede neural depende de um componente aleatório para a inicialização dos pesos e dessa forma não é afirmativamente possível obter-se os mesmos

Tempo	Observado	Predito.MLP	Diferença MLP
1	13,98	14,7430224	0,763022403
2	14	14,51348215	0,513482154
3	14,11	14,61757297	0,507572965
4	13,6	14,60230485	1,002304854
5	13,2	14,23767821	1,037678214
6	12,6	13,87146482	1,271464816
7	12,45	13,18193909	0,73193909
8	12,78	12,98831193	0,208311929
9	14,3	13,16877598	-1,131224024
10	14,21	14,55304605	0,343046045
11	14,15	14,75573728	0,605737279
12	14,1	15,02342809	0,923428093
13	13,44	14,65924339	1,219243394
14	12,8	14,09424573	1,294245731
15	12,32	13,48581068	1,16581068
16	12,13	12,87074597	0,740745967
17	12,73	12,6172981	-0,112701901
18	12,23	13,09099369	0,860993694
19	12,26	12,75667107	0,496671073
20	11,5	12,92883679	1,428836789
21	11,36	12,04598095	0,685980952
22	10,83	11,94760096	1,117600959

Figura 25: Cenário 2 - Experimento 15 - Dados da Previsão

resultados em duas rodadas distintas de processamento serão os mesmos. O software aqui utilizado, R-Studio que implementa a linguagem de programação R tem em seu conjunto de instruções a função `set.seed` que fixa a geração de números aleatórios para esses pesos, todavia, por vezes ela repete os mesmos valores para todas as entradas tornando os resultados repetidos, logo inválidos. Para o experimento em Random Walk não há essa componente aleatória uma vez que ele depende exclusivamente da série de dados de entrada e da saída desejada.

Seguindo com os experimentos, dessa feita com a melhor configuração dos ensaios anteriores para cada cenário tem-se na figura 28 os resumos dos resultados obtidos. É importante destacar o número de interações que levou a rede MLP à convergência em cada experimento, sabendo-se que foram estabelecidas 10.000 épocas para o processamento, ainda, a rede neural obteve esses resultados com os seguintes parâmetros:

a.) No cenário 1

Nos experimentos de 1 a 8, foram 5 neurônios e taxa de aprendizado igual a 0,8. Para os experimentos de 9 a 16 foram 10 neurônios e taxa de aprendizado igual a 0,8.

b.) No cenário 2

Nos experimentos de 1 a 8, foram 5 neurônios e taxa de aprendizado igual a 0,7. Para os experimentos de 9 a 16 foram 10 neurônios e taxa de aprendizado igual a 0,7. As figuras 26 e 27 mostram o comportamento dos resultados do processo preditivo na hipótese de Random Walk, percebe-se que em ambos os cenários os resultados da predição são muito próximos dos valores observados, o que recomenda o processo para predição.

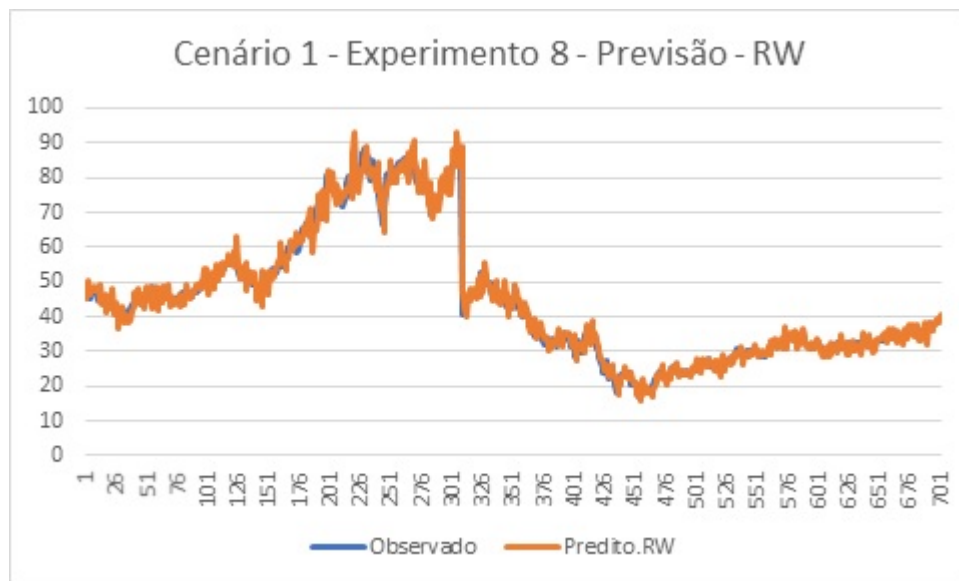


Figura 26: Cenário 1 - Experimento 8 - Previsão - RW (Random Walk)

Foram executadas 5 cinco rodadas de 16 execuções cada para a rede com 5 e 10 neurônios, obtendo-se os resultados acima. Para as execuções de cada conjunto de experimentos a amostra foi alterada aleatoriamente dentro dos parâmetros de 70% para treinamento e 30% para a previsão, dessa forma buscamos evitar eventuais condições que viesse a favorecer os resultados.

Da análise das médias dos erros obtidos tem-se que a Hipótese de Random Walk mostrou-se bem mais eficiente que a Rede Neural MLP, tanto no cenário 1 com 5 neurônios e taxa de aprendizagem de 0,8 com erro médio de 2,49% contra 4,99% da rede neural, como no cenário 2 com 10 neurônios e taxa de aprendizagem 0,7 onde obteve erro médio de 1,89% contra 7,58% da rede neural. Ainda, tomando-se por base de análise o desvio padrão tem-se resultados inferiores para a Hipótese de Random Walk, reforçando o bom

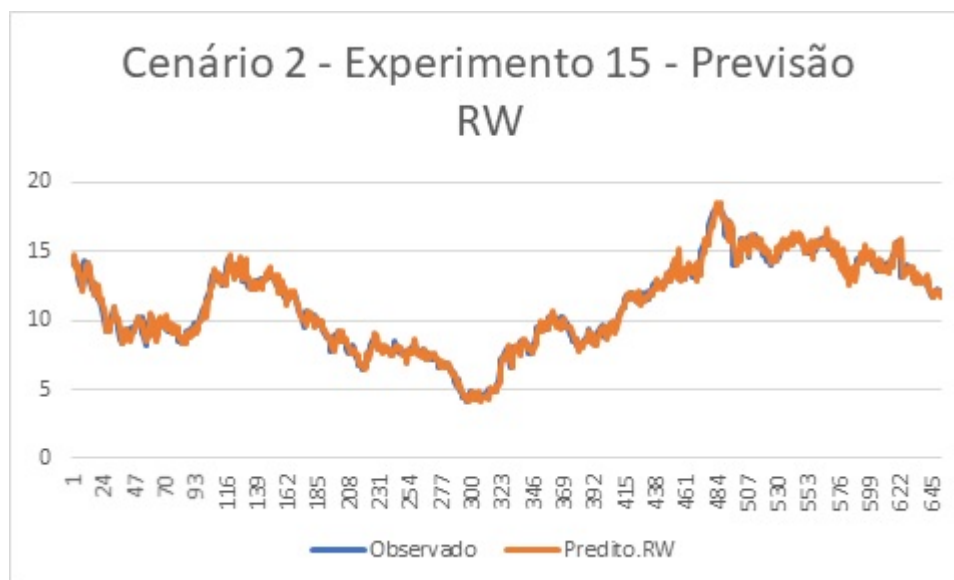


Figura 27: Cenário 2 - Experimento 15 - Previsão - RW (Random Walk)

Resultados						
Cenário 1				Cenário 2		
Experimento	Interações	Erro RMSE - MLP	Erro RMSE - RW	Interações	Erro RMSE - MLP	Erro RMSE - RW
1	1230	0,057659278	0,025535941	660	0,114504915	0,019249829
2	1760	0,026703787	0,02424997	1360	0,133119741	0,019165655
3	1370	0,051365495	0,025742461	870	0,163654717	0,018089722
4	880	0,035786861	0,024461559	970	0,037707915	0,019378211
5	750	0,033425514	0,025251587	950	0,116923002	0,019327209
6	1150	0,057466288	0,024541882	650	0,039374456	0,018482957
7	910	0,069608328	0,024651951	200	0,091490599	0,018824853
8	730	0,043099599	0,024588257	1410	0,032347052	0,018327835
9	1020	0,091900937	0,02372751	1170	0,04095351	0,018655488
10	2110	0,027339725	0,024984203	820	0,075487526	0,018555133
11	1840	0,028145167	0,026149424	1420	0,100892976	0,019149326
12	1300	0,0423	0,025489412	260	0,106325683	0,018525667
13	890	0,078071939	0,024814822	490	0,027002221	0,019122994
14	1840	0,059189989	0,024289655	1020	0,05988059	0,019128395
15	1340	0,058555567	0,025196787	750	0,039719557	0,019082994
16	1180	0,03745168	0,025054243	540	0,032788732	0,019627004
Média		0,049879385	0,024920604		0,075760825	0,018918329
Desvio Padrão		0,01910817	0,000627941		0,042947721	0,000432603
Percentual		4,99%	2,49%		7,58%	1,89%

Figura 28: Quadro comparativo dos resultados entre a rede neural MLP com a hipótese de Random Walk

desempenho do processo.

Analisando-se os resultados obtidos depreende-se que ambos os processos são demasiadamente bons para a previsão de séries temporais, acompanhando a não linearidade

do comportamento da série e trazendo resultados muito próximos do observado. Não se pode em razão das diferenças comparativas observadas excluir a rede neural desse processo preditivo, pois o resultado individual também é importante.

Salvaguardando a volatilidade dos preços das ações no mercado financeiro, que sofrem variações bruscas por vezes, esse processo de previsão, com rede neural ou hipótese de Random Walk, podem servir para orientações de investimentos, mas recomenda-se aqui bastante cautela para investimentos em horizontes de médio e longo prazo.

5 CONCLUSÕES

Atendendo a proposta inicial, esse trabalho mostrou que as redes neurais podem ser usadas para previsão em séries temporais, apresentando um resultado bastante favorável. O trabalho tomou como base de dados dois períodos críticos da economia, o primeiro com a Crise Financeira Mundial de 2008 e o segundo com o período de recessão econômica no Brasil. Nesses dois cenários foi possível verificar a volatilidade das ações no mercado financeiro e a não linearidade no comportamento da série de valores dos preços. Com os valores submetidos a processamento pela rede neural artificial MLP pode-se obter resultados previstos muito próximo dos valores observados, indicando dessa forma que as redes neurais podem ajudar os analistas financeiros de investimentos nas suas atividades.

É importante marcar que o processo aqui apresentado é apenas para orientação sobre o comportamento das ações e previsões no curto prazo, não devendo ser tido como algo certo e isolado para fins de investimentos. Quem opera o mercado financeiro sabe muito bem do comportamento e o quanto ele é variável em razão de aspectos políticos e sociais, que são as variáveis exógenas no processo e de difícil controle e previsão de ocorrência e do impacto que vai causar.

Buscar em estudos futuros relacionar aspectos sociais, políticos e comportamentais às variáveis econômicas, tais como PIB, Renda Percapita, Taxa de juros e outras, e havendo a devida correlação, isso pode refinar esse processo preditivo e melhorar os resultados obtidos, dando mais confiabilidade ao processo preditivo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADLER, A. S. e. a. *Árvores Binomias Implícitas: aplicação para as opções de Telebrás no exercício de abril de 1999*. 1. ed. Foz do Iguaçu: XXIII ENANPAD, 1999.
- AIT-SHALAIA Y., L. A. *Nonparametric estimation of state-price densities implicit in financial asset prices*. 1. ed. USA: University of Chicago and MIT, 1996.
- AMIM K., J. E. *Pricing options on risky assets in a stochastic interest rate economy*. v.2, p.217-237. USA: Mathematical Finance, 1992.
- AMIN K., N. V. *Option valuation with systematic stochastic volatility*. v. 48, p.881-910. USA: Journal of Finance, 1993.
- BAILEY W., S. R. *The pricing of stock index options in a general equilibrium model*. v.24, p.1-12. USA: Journal of Financial and Quantitative Analysis, 1993.
- BAKSHI G., C. C. C. Z. *Empirical performance of alternative option pricing models*. v.52, n.5, p.2003-2049. USA: Journal of Finance, 1997.
- BAKSHI G., C. C. C. Z. *Empirical performance of alternative option pricing models*. v.52, n.5, p.2003-2049. USA: Journal of Finance, 1997.
- BARROS, A. C. e. a. *Análise de Séries Temporais em R*. 1. ed. São Paulo: Editora Elsevier, 2018.
- BATES, D. *The crash of 87: was it expected? The evidence from options markets*. v. 46, p.1009-1044. USA: Journal of Finance, 1991.
- BATES, D. *Testing option pricing models*. vol.15: Statistical methods in finance, p.567-611. North Holland, Amsterdam: Journal of Financial and Quantitative Analysis, 1996.
- BATTITI, R. *First and Second Order Methods for Learning: Between Steepest Descent and Newton's Method*. 1. ed. London: Neural Computation, v.4. no.2, p. 141-166, 1992.
- BONOMO, M. *Finanças aplicadas ao Brasil*. 1. ed. Rio de Janeiro: FGV, 2002.

- BRITO NEY O.; MANAZES, J. C. F. *A eficiência informacional fraca do mercado à vista na Bolsa de Valores do Rio de Janeiro no período 1973-1980*. 1. ed. Rio de Janeiro: Relatório de Pesquisa, 1981.
- BRONSTEIN G., F. H. G. M. E. S. S. L. E. F. *Modelo de avaliação de opções para processos de difusão e salto*. v. 39, p. 147-166. Foz do Iguaçu: XXIII ENANPAD, 1999.
- BUENO, R. d. L. d. S. *Econometria de Séries Temporais*. 2. ed. São Paulo: Cengage Learning, 2012.
- CAMPBELL, J. Y. e. a. *The Econometrics of Financial Markets*. 1. ed. Princeton, N.J.: Princeton University Press, 1997.
- CARTACHO, M. S. *A utilização de um modelo híbrido Algoritmo Genético/Redes Neurais no processo de seleção de carteiras*. Dissertação (Mestrado) — Dissertação (Mestrado em Administração) – Faculdade de Ciências Econômicas, UFMG, Belo Horizonte, 2001.
- CASTRO, C. L. d. *Sistema Inteligente para Previsão on-line do Índice Ibovespa baseado em Redes Neurais Artificiais*. Dissertação (Mestrado) — Dissertação (Mestrado em Administração) – Faculdade de Ciências Econômicas, UFMG, Belo Horizonte, 2003.
- CHIRAS, D. P. *The Information content of option prices and a test of market efficiency*. 1. ed. USA: Journal of Finance., 1978.
- CHRYSSOLOURIS, G. M. L. e. A. R. *Confidence interval prediction for neural network models*. Vol. 7, no. 1, pg.229-232. EUA: IEEE Transactions on Neural Networks, 1996.
- COX J., R. S. *The valuation of options for alternative stochastic processes*. v.3, p.145-166. USA: Journal of Financial Economics, 1976.
- COX J., S. R. M. R. *Option pricing: a simplified approach*. n.7, p.229-263. USA: Journal of Financial Economics., 1979.
- DAIGLER, R. T. *Advanced options trading: the analysis and evaluation of trading strategies, hedging tactics & pricing models*. 1. ed. EUA: McGraw-Hill, 1997.
- DERMAN E., K. I. R. *Riding on a smile*. v.7, n.2, p.32-39. USA: Risk, 1994.

- DUPIRE, B. *Pricing with a smile*. v.7, n.1,p. 18-20. USA: Risk, 1994.
- FAMA, E. F. Efficient capital markets: a review of theory and empirical work. *Journal of Finance: 1970, maio, pp. 383-417.*, 1970.
- FREITAS, S. O. *Utilização de um modelo baseado em Redes Neurais para a precificação de opções*. Dissertação (Mestrado) — Dissertação (Mestrado em Administração) – Faculdade de Ciências Econômicas, UFMG, Belo Horizonte, 2001.
- GEMAN, S. E. B. e. R. D. *Neural networks and the bias/variance dilemma*. Vol. 4, pg. 1-58. EUA: Neural Computation, 1992.
- GESKE, R. *The valuation of compound options*. n. 7, p. 63-81. USA: Journal of Financial Economics., 1979.
- HAGAN, M. T. e. a. *Neural Network Design*. 1. ed. Boston, MA: PWS, 1996.
- HAYKIN, S. *Neural networks and learning machines*. 2. ed. London: Pearson Education, 1999.
- HERTZ, J. e. a. *Introduction to the theory of neural computation*. 1. ed. Redwood City: Addison-Wesley, 1991.
- HESTON, S. *A closed-form solution for options with stochastic volatility with applications to bond and currency options*. v.6, p.327-343. USA: Review of Financial Studies, 1993.
- HULL, J. *Introdução aos mercados futuros e de opções*. 1. ed. São Paulo: BM&F, 1996.
- HULL, J. *Opções, futuros e outros derivativos*. n. 7, p. 63-81. São Paulo: BM&F, 1998.
- HULL, J. C. *The pricing of options on assets with stochastic volatilities*. 1. ed. USA: Journal of Finance, 1987.
- JARROW R, R. A. *Approximate option valuation for arbitrary stochastic processes*. v.10, p.347-370. USA: Journal of Finance, 1982.
- KON, S. J. *Models of stock returns – a comparison*. v. 39, p. 147-166. USA: Journal of Finance, 1984.

- LANARI, C. S. *O efeito "sorriso" da volatilidade implícita do modelo de Black e Scholes: estudo empírico sobre as opções Telebrás PN no ano de 1998*. 1. ed. Belo Horizonte: Dissertação de Mestrado. CEPEAD/UFMG., 2000.
- MADAN D., C. E. *The VG option pricing model, Working paper*. 1. ed. USA: University of Maryland and Georgia Institute of Technology, 1996.
- MELINO A., T. S. *Pricing foreign currency options with stochastic volatility*. v.45, p.239-265. USA: Journal of Econometrics, 1990.
- MELINO A., T. S. *Misspecification and the pricing and hedging of long-term foreign currency options*. v.14, p.373-393. USA: Journal of International Money and Finance, 1995.
- MENDES, L. S. *Avaliação de opções: uma comparação do desempenho de métodos distintos de se estimar a volatilidade*. 1. ed. Belo Horizonte: Dissertação de Mestrado. CEPEAD/UFMG., 2000.
- MERTON, R. C. *Theory of rational option pricing*. n. 4, p. 141-183. USA: Bell Journal of Economics and Management Science., 1973.
- MERTON, R. C. *Option pricing when underlying stock returns are discontinuous*. n. 3, p.125-144. USA: Journal of Financial Economics., 1976.
- MILLER, G. F. e. a. Designing neural networks using genetic algorithms. *Third International Conference on Genetic Algorithms. Anais, p. 177-179. San Mateo: Morgan Kaufmann*, 1989.
- NETO, L. A. S. *Opções: do tradicional ao exótico*. 1. ed. São Paulo: Atlas, 1996.
- PARRA DOMINGOS FILHO; SANTOS, J. A. *Metodologia Científica*. 5. ed. São Paulo: Ed. Futura, 2002.
- REFENES, A. N. e. a. *Financial Modelling using Neural Networks*. 1. ed. London: Department of Computer Sciences, University College London, 1993.
- RENDLEMAN, R. J. *Standard deviations of stock price ratios implied in options prices*. 1. ed. USA: Journal of Finance., 1976.

- RESENDE, F. S. d. *Ciência de dados aplicado à BM&FBovespa*. 1. ed. Curitiba: Appris Editora, 2016.
- RUBINSTEIN, M. *Implied binomial trees*. v. 49, n. 3, p. 771-818. USA: Journal of Finance., 1994.
- RUMELHART, D. E.; MCCLELLAND, J. L. *Parallel Distributed Processing: explorations in the microstructure of cognition, vols. 1 e 2*. 4. ed. Cambridge, MA, EUA: The MIT Press, 1986.
- RUSSEL STUART J.; NORVIG, P. *Inteligência Artificial*. Rio de Janeiro: Elsevier, 2003.
- SAMUELSON, P. *Proof that Properly Anticipated Prices Fluctuate Randomly*. 1. ed. Chicago-USA: Industrial Management Review, 1965.
- SANTOS, A. R. d. *Metodologia Científica*. 5. ed. Rio de Janeiro: Ed. DP e A, 2002.
- SCOTT, L. O. *Option pricing when the variance changes randomly: theory, estimation, and an application*. vol. 22, p. 419-438. USA: Journal of Financial and Quantitative Analysis, 1987.
- SCOTT, L. O. *Option pricing when the variance changes randomly: theory, estimation, and an application*. vol. 22, p. 419-438. USA: Journal of Financial and Quantitative Analysis, 1987.
- SILVA, I. N. d. *Redes neurais artificiais: para engenharia e ciências aplicadas*. 2. ed. São Paulo: Artliber Editora Ltda, 2016.
- SILVA, L. A. D. e. a. *Introdução a Mineração de Dados com Aplicações em R*. 1. ed. São Paulo: Elsevier, 2016.
- STEIN E., S. J. *Stock price distributions with stochastic volatility*. v.4, p.727-752. USA: Review of Financial Studies, 1991.
- WIGGINS, J. B. *Option values under stochastic volatility: theory and empirical estimates*. Vol. 19, 1987, p.351-372. USA: Journal of Financial and Quantitative Analysis, 1987.

WOOLDRIDGE, J. M. *Introdução a Econometria – Uma abordagem moderna*. 4. ed.
São Paulo: Cengage Learning, 2016.