

**Instituto de Pesquisas Tecnológicas do Estado de São Paulo**

**Willian Oliveira Ribeiro**

**Estudo sobre a empregabilidade de variáveis  
macroeconômicas no uso de redes neurais para a previsão  
do preço de fechamento no mercado de ações brasileiro**

**São Paulo  
2019**

Willian Oliveira Ribeiro

Estudo sobre a empregabilidade de variáveis macroeconômicas no uso de redes neurais para a previsão do preço de fechamento no mercado de ações brasileiro

Dissertação de Mestrado apresentada ao Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT, como parte dos requisitos para a obtenção do título de Mestre em Engenharia de Computação.

Data da aprovação \_\_\_\_/\_\_\_\_/\_\_\_\_

---

Prof. Dr. Fábio Silva Lopes (Orientador)  
Mestrado Engenharia de Computação

Membros da Banca Examinadora:

Prof. Dr. Fábio Silva Lopes (Orientador)  
Mestrado Engenharia de Computação

Prof. Dr. Marcelo Novaes de Rezende (Membro)  
Mestrado Engenharia de Computação

Prof. Dr. Mario Olimpio de Menezes (Membro)  
IPEN – Instituto de Pesquisas Energéticas Nucleares

Willian Oliveira Ribeiro

Estudo sobre a empregabilidade de variáveis macroeconômicas no uso de redes neurais para a previsão do preço de fechamento no mercado de ações brasileiro

Dissertação de Mestrado apresentada ao Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT, como parte dos requisitos para a obtenção do título de Mestre em Engenharia da Computação: Engenharia de Software.

Área de Concentração: Engenharia de Software

Orientador: Prof. Dr. Fábio Silva Lopes

São Paulo  
Julho/2019

Ficha Catalográfica  
Elaborada pelo Departamento de Acervo e Informação Tecnológica – DAIT  
do Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT

R484e **Ribeiro, Willian Oliveira**

Estudo sobre a empregabilidade de variáveis macroeconômicas no uso de redes neurais para a previsão do preço de fechamento no mercado de ações brasileiro. / Willian Oliveira Ribeiro. São Paulo, 2019.  
83p.

Dissertação (Mestrado em Engenharia de Computação) - Instituto de Pesquisas Tecnológicas do Estado de São Paulo. Área de concentração: Engenharia de Software.

Orientador: Prof. Dr. Fábio Silva Lopes

1. Redes neurais artificiais 2. Mercado de ações 3. Algoritmo de aprendizado 4. Tese I. Lopes, Fábio Silva, orient. II. IPT. Coordenadoria de Ensino Tecnológico III. Título

19-51

CDU 004.032.26(043)

## **RESUMO**

O interesse em usar redes neurais artificiais (RNAs) para previsão levou a um grande aumento nas atividades de pesquisa nas últimas décadas, embora as RNAs forneçam uma grande promessa com resultados comprovados, elas também incorporam muita incerteza. Esta situação também é observada no mercado financeiro de ações com a previsão de preços de ações usando técnicas de aprendizado de máquina, em especial, redes neurais aplicadas que têm sido amplamente pesquisadas e utilizadas. Neste contexto torna-se importante entender o algoritmo e as informações que embasam seu processo decisório. O presente estudo tem por objetivo avaliar o uso de redes neurais artificiais, o algoritmo de aprendizado e a influência de fatores macroeconômicas no mercado de ações brasileiro. A avaliação da rede dar-se-á pela métrica RMSE.

Palavras Chaves: Previsão de ações; redes neurais artificiais; fatores macroeconômicos

## **ABSTRACT**

### **Study on the employability of macroeconomic variables in the use of neural networks for forecasting the closing price in the stock market**

The interest in using artificial neural networks (ANNs) for prediction has led to a large increase in research activities in the last decades, although ANNs provide great promise with proven results, they also incorporate a great deal of uncertainty. This situation is also observed in the financial stock market with the forecast of stock prices using machine learning techniques, in particular, applied neural networks that have been widely researched and used. In this context it becomes important to understand the algorithm and the information that underlies its decision-making process. The present study aims to evaluate the use of artificial neural networks and the influence of macroeconomic factors on the Brazilian stock market. The evaluation of the network will be by RMSE metrics.

**Keywords:** Stock market prediction, artificial neural networks, macroeconomic factors

## Lista de ilustrações

<b>Figura 1 –</b>	Volume financeiro negociado na Bolsa de Valores Brasileira (BM&F Bovespa e B3) entre 2009 e 2018	22
<b>Figura 2 –</b>	Representação entre oferta e demanda de ações	24
<b>Figura 3 –</b>	Relação de volume financeiro e evolução do preço das ações ITUB4, VALE3 e BBDC4	25
<b>Figura 4 –</b>	Exemplo de linhas de tendência da ação PETR4	29
<b>Figura 5 –</b>	Representação de uma rede Hopfield	36
<b>Figura 6 –</b>	Representação de um neurônio biológico	37
<b>Figura 7 –</b>	Representação matemática de um neurônio do tipo Perceptron	38
<b>Figura 8 –</b>	Funções de ativação: Linear, Degrau, Sigmóide e Tangente Hiperbólica	41
<b>Figura 9 –</b>	Representação de rede neural multilayer perceptron	42
<b>Figura 10 –</b>	Representação de uma célula de uma rede LSTM	44
<b>Figura 11 –</b>	Representação de uma rede LSTM composta por 3 células	45
<b>Figura 12 –</b>	Diagrama do processo experimental a ser aplicado	57
<b>Figura 13 –</b>	Representação da evolução do preço médio do ETF BOVA11 no período de 03/01/2008 a 31/12/2018	58
<b>Figura 14 –</b>	Representação de uma rede neural artificial com uma linha de atraso	66
<b>Figura 15 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BOVA11	70
<b>Figura 16 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo VALE3	71
<b>Figura 17 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM o ativo PETR4	72

<b>Figura 18 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo ITUB4	73
<b>Figura 19 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BBDC4	74
<b>Figura 20 –</b>	Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BBSA3	75



## Lista de tabelas

<b>Tabela 1 –</b>	Subconjunto da amostra com os 5 primeiros registros do arquivo – BOVA11	57
<b>Tabela 2 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo BOVA11	61
<b>Tabela 3 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo VALE3	62
<b>Tabela 4 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo PETR4	63
<b>Tabela 5 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo ITUB4	63
<b>Tabela 6 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo BBDC4	64
<b>Tabela 7 –</b>	Demonstração da correlação (Pearson, Spearman e Kendall) com a variável alvo, Preço de Fechamento – Ativo BBSA3	65
<b>Tabela 8 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BOVA11	70
<b>Tabela 9 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo VALE3	71
<b>Tabela 10 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo PETR4	72
<b>Tabela 11 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo ITUB4	73
<b>Tabela 12 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BBDC4	74
<b>Tabela 13 –</b>	Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BBSA3	75
<b>Tabela 14 –</b>	Resultados consolidados para os papeis avaliados	76

## **Lista de quadros**

<b>Quadro 1 –</b>	Trabalhos relacionados	48
<b>Quadro 2 –</b>	Comparativo entre a Design Science Research, o Estudo de Caso e a Pesquisa-Ação	56
<b>Quadro 3 –</b>	Descrição das variáveis técnicas independentes	59
<b>Quadro 4 –</b>	Descrição das variáveis contábeis e macroeconômicas	60

## Lista de abreviaturas e siglas

ADALINE	Neurônio Linear Adaptativo (Adaptive Linear Neuron).
ARIMA	Modelo auto-regressivo integrado de médias móveis.
B3	Bolsa de mercado de capitais correspondente a fusão das empresas CETIP e BMF&BOVESPA.
BA	Aprendizado Bayesiano ou Inferência Bayesiana.
BBDC4	Ação preferencial do Banco Bradesco.
BBSA3	Ação ordinária do Banco do Brasil.
BIST	Mercado de ações de Istambul.
BRICS	Acrônimo que se refere aos países: Brasil, Rússia, Índia, China e África do Sul que juntos formam um time de cooperação econômica e política.
BOVA11	ETF comercializado na B3 que reflete o Ibovespa.
CG	Método do gradiente conjugado.
CL	Método de aprendizado competitivo.
GD	Método do gradiente descendente.
HE	Método de aprendizagem hebbiana.
HY	Método de aprendizado híbrido.
IBOVESPA	Índice da Bolsa de Valores de São Paulo, é um indicador de desempenho das ações negociadas na B3. É formado pelas ações com maior volume negociado nos últimos meses.
IPO	Initial Public Offer ou Oferta Pública Inicial.
ITUB4	Ação preferencial do Itaú Unibanco.
LPA	Lucro por ação.
LSTM	Long Short Term Memory.
MAPE	Mean Absolute Percentage Error.
MLP	Multilayer Perceptron.

PETR4	Ação preferencial da Petrobrás.
QN	Método Quasi-Newton.
RMSE	Root Mean Square Error.
RNA	Rede Neural Artificial.
RPL	Retorno sobre patrimônio líquido.
SSE	Soma dos erros ao quadrado.
VALE3	Ação ordinária da Vale do Rio Doce.
VLSI	Very Large Scale Integrated.

## SUMÁRIO

1 INTRODUÇÃO	15
2 OBJETIVOS	19
2.1 Objetivo	19
2.2 Contribuições	19
2.3 Organização do trabalho	19
3 REVISÃO DA LITERATURA	21
3.1 Introdução	21
3.2 Sistema financeiro nacional	21
3.3 Macroeconomia	21
3.4 Mercado de ações	22
3.5 Ibovespa	24
3.6 Análise fundamentalista	26
3.7 Análise técnica	27
3.8 Séries temporais	29
3.9 Redes neurais artificiais	31
3.9.1 Contexto histórico	34
3.9.2 Estrutura do neurônio	36
3.9.3 Características de redes neurais artificiais	39
3.9.4 Função de ativação	39
3.9.5 Arquitetura de redes neurais	41
3.9.5.1 Rede multilayer perceptron	41
3.9.5.2 Redes LSTM	43
3.10 O processo de aprendizagem de redes neurais artificiais	45
3.11 O algoritmo de retropropagação e outros algoritmos de treinamento	46
3.12 Trabalhos relacionados	48

<b>4 MÉTODO DE PESQUISA</b>	<b>54</b>
4.1 Introdução	54
4.2 Método de Pesquisa	54
4.3 Configurações experimentais	55
4.4 Conjunto de dados	56
4.4 Janelas móveis	65
4.5 Python	66
4.6 Arquitetura das Redes Neurais e configurações experimentais	66
4.7 RMSE	68
<b>5 ANÁLISE DOS RESULTADOS</b>	<b>69</b>
<b>6 CONCLUSÕES E RECOMENDAÇÕES</b>	<b>77</b>
<b>REFERÊNCIAS</b>	<b>79</b>

## 1 INTRODUÇÃO

O Mercado de ações está diretamente relacionado com a economia de um país. Quanto mais desenvolvida é uma economia, mais ativo é seu mercado de capitais, o que se traduz em mais oportunidades para as pessoas, empresas e instituições aplicarem suas poupanças (Silva, 2015). A Bolsa de valores é um dos componentes do mercado de ações e tem por objetivo principal prover meios necessários que possibilitem a comercialização de produtos diversos, por exemplo, ações de empresas que representam uma fração do capital social que quando adquirida pelo investidor torna-o sócio. O objetivo da empresa que vende papeis na bolsa de valores é a captação de recursos e o do investidor é o desenvolvimento de seu patrimônio.

A inconstância no preço de compra e venda de ações é uma característica que possibilita aos investidores a possibilidade de lucro e também o prejuízo sobre o capital investido. Esta variação pode ser influenciada por diversos fatores, por exemplo: O preço do ouro comercializado, o preço do barril do petróleo, a situação econômica do país em que a bolsa opera, a situação econômica de outros governos, variação cambial, taxas oficiais, regulamentações, abertura de capital de outras empresas na bolsa e muitos outros fatores econômicos. Neste cenário os investidores acompanham a variação do mercado para que suas escolhas sejam mais precisas e a possibilidade de lucro seja maior. Esta situação é dificultada pela grande quantidade de informações as quais temos acesso atualmente.

Identificar os fatores importantes para explicar os preços de ações contemporâneos tem sido um foco da literatura acadêmica, com a pesquisa divisível em dois ramos principais, distintos e muitas vezes concorrentes: Análise fundamental e técnica (Bettman et al., 2009).

Segundo Bettman et al. (2009), um dos primeiros a formalizar uma argumentação sobre a importância dos fatores fundamentalistas nos exercícios de valorização das ações foi Graham e Dodd (1934). Desde então diversos estudos evoluíram com o intuito de aperfeiçoar esta técnica que é utilizada até hoje. Na análise fundamentalista o valor intrínseco da ação é mensurado de acordo com outros fatores externos a ele, por exemplo, opinião pública sobre a empresa, balanço comercial, corpo diretivo e menções na mídia. Fatores qualitativos são transformados em quantitativos para comparação com o preço corrente da ação.

Complementar à análise fundamentalista existe a análise técnica em que o preço passado de ações é utilizado para a análise de potencial lucro de uma ação baseado em seu preço corrente e histórico. O reconhecimento do potencial de preços passados e os movimentos nele previstos para prever valores de patrimônio futuros é atribuído primeiramente a uma série de editoriais publicados por Charles Dow no Wall Street Journal entre 1900 e 1902 (Bettman et al., 2009).

Os métodos de análise fundamentalista e técnica podem subsidiar a tomada de decisão do investidor e são utilizadas com aceitação no mercado de investimentos. Este tipo de análise foi e ainda é alvo de inúmeros estudos da área financeira e recentemente de estudos da área da computação. Muitos investidores fazem uso de informações financeiras para monitorar a situação do mercado, planejar investimentos e antecipar riscos.

Dentre os muitos estudos na área econômica derivados da análise destacam-se as hipóteses do mercado eficiente e a de caminhada aleatória. Apesar da influência e aceitação das hipóteses, são apenas teorias e uma teoria pode ser refutada com evidências apropriadas.

Segundo Fama (1964) a hipótese do mercado eficiente afirma que o preço de uma ação reflete a assimilação de toda a informação correspondente naquele momento, o preço é corrigido a medida que novas informações são divulgadas, portanto dados atuais ou históricos não são significativos para previsões ou mudanças futuras de preços.

Também segundo Fama (1965) a hipótese de caminhada aleatória afirma que os preços das ações não dependem dos preços passados. Assim, estes não são padrões a serem explorados, os dados históricos não refletem o padrão do preço atual das ações, mas sim o preço correspondente ao período analisado.

White (1988) contestou a teoria do mercado eficiente com a aplicação de redes neurais artificiais para previsão do preço de fechamento de ações da IBM. Apesar do fracasso do experimento, as sugestões do pesquisador de uma rede neural artificial que permitisse o uso de mais dados de entrada (por exemplo, volume, o preço de outras ações, volume, dados macroeconômicos) e o uso de diferentes horizontes de tempo é adotado ao longo de tempo em diversos estudos.



Atsalakis e Valavanis (2009) concluíram em seu estudo que as redes neurais estão entre as técnicas mais utilizadas para prever valores futuros ou tendências dos mercados de ações ao redor do mundo.

Técnicas para melhorar a acurácia da previsão e aumentar as possibilidades de lucro sobre o investimento realizado foram desenvolvidas e evoluídas embasadas em modelos estatísticos e também na experiência dos investidores. A integração com a tecnologia fez com que muitas técnicas fossem automatizadas e que diversas variáveis fossem testadas para a decisão de investimento mais precisa possível.

Cabe ao investidor encontrar formas para maximizar seus lucros de acordo com as informações utilizadas para a tomada de decisão. Por exemplo, o experimento desenvolvido por Khare et al.(2009) através da comparação de dois tipos de arquitetura de redes neurais LSTM e MLP para a previsão do preço de fechamento de ações em um curto espaço de tempo mostra a importância da arquitetura da rede neural a ser utilizada e das variáveis de entrada. Um outro ponto de atenção é a evidência do experimento contra a teoria do mercado eficiente.

Por se tratar de um problema caracterizado por uma série temporal, as redes neurais recorrentes e profundas podem ser mais apropriadas para este tipo de problema do que outros algoritmos em que não haja memória de curto prazo (Graves, 2012).

Tkáč e Verner (2016) realizaram uma revisão sistemática para identificar e analisar publicações correspondentes ao uso de redes neurais no ambiente corporativo entre 1994 e 2015, concluindo que os assuntos mais abordados para a aplicação de redes neurais foram: Análise de dificuldades financeiras e falências, classificação de crédito e previsão do preço de ações. Na maioria dos artigos o modelo de rede neural proposto obteve melhores resultados quando comparado a outras técnicas como análise discriminante e regressão linear.

Nazário et al. (2017) avaliaram artigos sobre análise técnica no mercado de ações. Foi verificado que 53 das pesquisas tinham como objeto de estudo países considerados de economia desenvolvida ou avançada, 23 relacionados a economias de países em desenvolvimento e 11 não puderam ser classificados. Concluiu-se que a literatura acadêmica correspondente a análise técnica em mercados emergentes, em especial sobre os BRICS, precisa ser explorada e desenvolvida.

Diante deste cenário os estudos sobre o tema indicam a necessidade de um aprofundamento maior em pesquisas sobre redes neurais e sobre o mercado financeiro de países emergentes, remanesce então a necessidade de detalhar estes objetos de estudo que são alvo desta pesquisa.

## 2 OBJETIVOS

A revisão sistemática de Tkáč e Verner (2016) mostra que apesar de existir muitos estudos sobre a aplicação de redes neurais no mercado financeiro o impacto de fatores macroeconômicos recebeu pouca atenção das pesquisas acadêmicas. Motivado por este problema, e com base em interesse pessoal e profissional do pesquisador, este estudo tem por objetivo responder o seguinte problema de pesquisa:

Qual abordagem de redes neurais que utiliza dados macroeconômicos e técnicos é a mais eficiente para realizar a previsão do preço de fechamento no mercado de ações brasileiro?

### 2.1 Objetivo

Em função da questão apresentada o objetivo desta pesquisa é o desenvolvimento de uma abordagem de redes neurais que contemple a utilização de dados macroeconômicos para avaliar o impacto destas variáveis e a previsibilidade do mercado de ações brasileiro. A avaliação dos resultados através das métricas .

### 2.2 Contribuições

A relevância do trabalho é expressa na contribuição para o interesse no uso de redes neurais artificiais no mercado de ações. As contribuições são as seguintes:

- Atualização de pesquisas anteriores sobre o uso de redes neurais no mercado de ações, para revalidar, contestar ou melhorar métodos e resultados.
- A pesquisa permite o conhecimento sobre um algoritmo com potencial sobre uma modalidade de investimento importante no Brasil.
- A presente dissertação contribuirá para a melhoria do uso de redes neurais para previsões na bolsa de valores B3, em especial sobre a aplicação de redes neurais e o uso de variáveis macroeconômicas.

### 2.3 Organização do trabalho

O capítulo 3, Revisão da Literatura descreve uma visão geral sobre redes neurais e sua aplicação no mercado financeiro e como os estudos publicados tem tratado o assunto, técnicas e resultados. Neste capítulo também é apresentado o

resultado de uma pesquisa sobre as técnicas existentes que utilizam fatores macroeconômicos para aperfeiçoar o resultado da previsão do algoritmo além da forma como este conhecimento será aplicado ao estudo em desenvolvimento.

O capítulo 4, Método de Pesquisa, detalha o trabalho desta pesquisa, seus respectivos passos os detalhes das atividades necessárias, as coletas que realizadas. Os resultados do experimento são apresentados no capítulo 5, Análise dos Resultados. A discussão dos resultados e a sugestão de trabalhos futuros é discutida no capítulo 6, Conclusões e Recomendações.

### 3 REVISÃO DA LITERATURA

#### 3.1 Introdução

Neste capítulo, serão apresentados: A fundamentação teórica que descreve o contexto do mercado de ações brasileiro, o conceito de redes neurais e suas principais propriedades e o estado da arte sobre o impacto de fatores macroeconômicos e redes neurais para a previsão do preço de ações, e as principais referências já publicadas que serão utilizadas para o desenvolvimento deste trabalho.

#### 3.2 Sistema financeiro nacional

O sistema financeiro nacional brasileiro é o conjunto de instituições e instrumentos que viabilizam o fluxo financeiro entre os poupadores e os tomadores na economia, segmentando-se entre o mercado monetário, mercado de crédito, mercado de câmbio e mercado de capitais (CVM, 2014). Mellagi e Ishikawa (2000) dizem que o mercado financeiro é o meio onde se realizam as transferências de recursos, ou fluxo dos recursos monetários, entre os agentes financeiros/econômicos de forma direta ou indireta.

A instituição que regulamenta o mercado de investimentos é a Comissão de Valores Mobiliários (CVM). Mellagi e Ishikawa (2000) definem a CVM como a entidade responsável pela regulamentação do mercado, trata-se de uma autarquia federal vinculada ao Ministério da Fazenda, instituída pela Lei no 6.385, de 7-12-1976, cujas atribuições são a normatização, regulamentação, desenvolvimento, controle e fiscalização do mercado de valores mobiliários do país. Devido a regulamentação da CVM de 2013, as instituições financeiras do Brasil tiveram de implementar mecanismos que garantissem a conformidade entre os interesses do cliente investidor, seus perfis de investimento e as aplicações que efetuassem.

#### 3.3 Macroeconomia

A macroeconomia, segundo Garcia e Vasconcellos (2002) estuda a economia como um todo, analisando a determinação o comportamento de grandes agregados, tais como: Renda e produtos nacionais, nível geral de preços, emprego e desemprego, estoque de moeda e taxas de juros, balança de pagamentos e taxa de câmbio. Sua estrutura pode ser dividida em 5 mercados que são:

- Mercado de bens e serviços

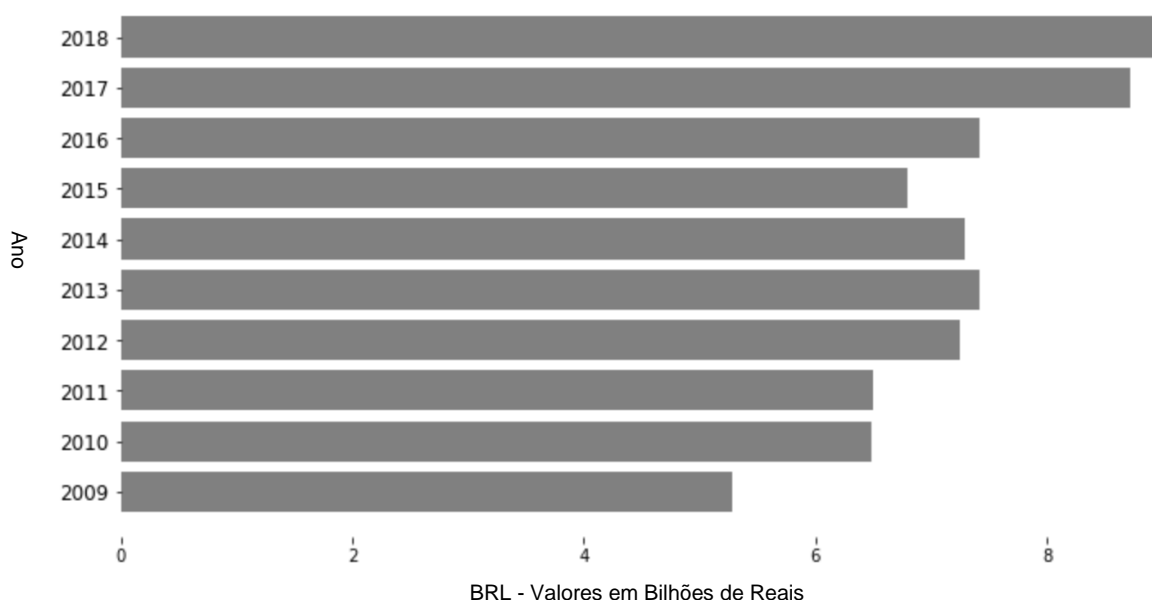
- Mercado de trabalho
- Mercado Monetário
- Mercado de Títulos/Ações
- Mercado de Divisas

No contexto brasileiro o contexto macroeconômico é objeto de estudo de diversas instituições públicas, por exemplo: IBGE, Banco Central. E particulares: FGV e B3.

### 3.4 Mercado de ações

O mercado de ações pode ser dividido em duas etapas: O mercado primário diz respeito à oferta inicial de ações enquanto que no mercado secundário ações já emitidas são comercializadas através da bolsa de valores. As bolsas de valores são o local criado e organizado pelas corretoras para negociação de valores mobiliários (Fortuna, 2005). Atualmente no Brasil existe somente uma bolsa de valores em operação, a B3, que é resultado da fusão da BM&F BOVESPA S.A. (Bolsa de Valores, Mercadorias e Futuros) e Cetip S.A. (Mercados Organizados) em março de 2017. A figura 1 demonstra o volume financeiro negociado na B3 nos últimos 10 anos.

**Figura 1** - Volume financeiro negociado na Bolsa de Valores Brasileira (BM&F Bovespa e B3) entre 2009 e 2018



**Fonte:** Elaborado pelo autor

A B3 (2018) define o conceito de ações como valores mobiliários emitidos por sociedades anônimas representativos de uma parcela do seu capital social. Em outras

palavras, são títulos de propriedade que conferem a seus detentores (investidores) a participação na sociedade da empresa. Elas são emitidas por empresas que desejam principalmente captar recursos para desenvolver projetos que viabilizem o seu crescimento. Uma ação representa uma parcela do capital social de uma empresa e ao comprar uma ação o investidor torna-se sócio daquela empresa. Uma oferta pública ou IPO é o primeiro passo para que a empresa comercialize ações que então podem ser compradas e vendidas na bolsa de valores. Depois do IPO e do início da comercialização das ações na bolsa de valores, o preço de uma ação pode ser definido através da relação entre: A quantidade de ações ofertadas que é sempre a mesma e a demanda da compra desta ação.

As ações são classificadas basicamente entre dois tipos: ordinárias (ON) e preferencias (PN). Os tipos PN, têm prioridade no recebimento de dividendos (antes dos acionistas ordinários) e os tipos ON, possuem como principal característica o direito de voto, ou seja, a capacidade de influenciar nas decisões da empresa (Assaf Neto, 2001).

As negociações de ações no mercado brasileiro são organizadas pela B3 que é uma empresa que provê serviços de negociação e pós-negociação (compensação e liquidação) de ações de empresas, derivativos de ações, financeiros e de mercadorias, títulos de renda fixa, títulos públicos federais, moedas à vista e commodities agropecuárias.

Segundo Fonseca (2013) o tema central para investir no mercado de ações, do ponto de vista da precificação de ativos consiste em vender ativos sobrevalorizados e comprar ativos subvalorizados. No entanto, para afirmar que o preço de um ativo está elevado ou baixo é preciso conhecer o valor do ativo em termos absolutos.

O preço pelo qual uma ação é negociada em bolsa, é fruto das condições de mercado (oferta e demanda) que refletem as condições estruturais e comportamentais da economia do país e específicas da empresa e de seu setor econômico (Fortuna, 2011). Segundo a B3 (2018), o preço da ação é formado pelo fluxo de oferta e procura de cada papel. Gráficamente a situação pode ser representada conforme a figura 2 onde o eixo vertical representa o preço da ação correspondente e o eixo horizontal a quantidade de ações. O ponto de intersecção entre a quantidade disponível (oferta) e a curva (demanda) pode ser entendido como o preço atual da ação. Este, é sujeito a oscilações, por exemplo: Publicação de balanços contábeis, divulgação dos lucros,

fusões, aquisições, mudança em fatores macroeconômicos, desenvolvimento de novos produtos e outros fatores. Acompanhar e prever estas oscilações pode oferecer oportunidades de lucro para investidores.

**Figura 2** - Representação entre oferta e demanda de ações



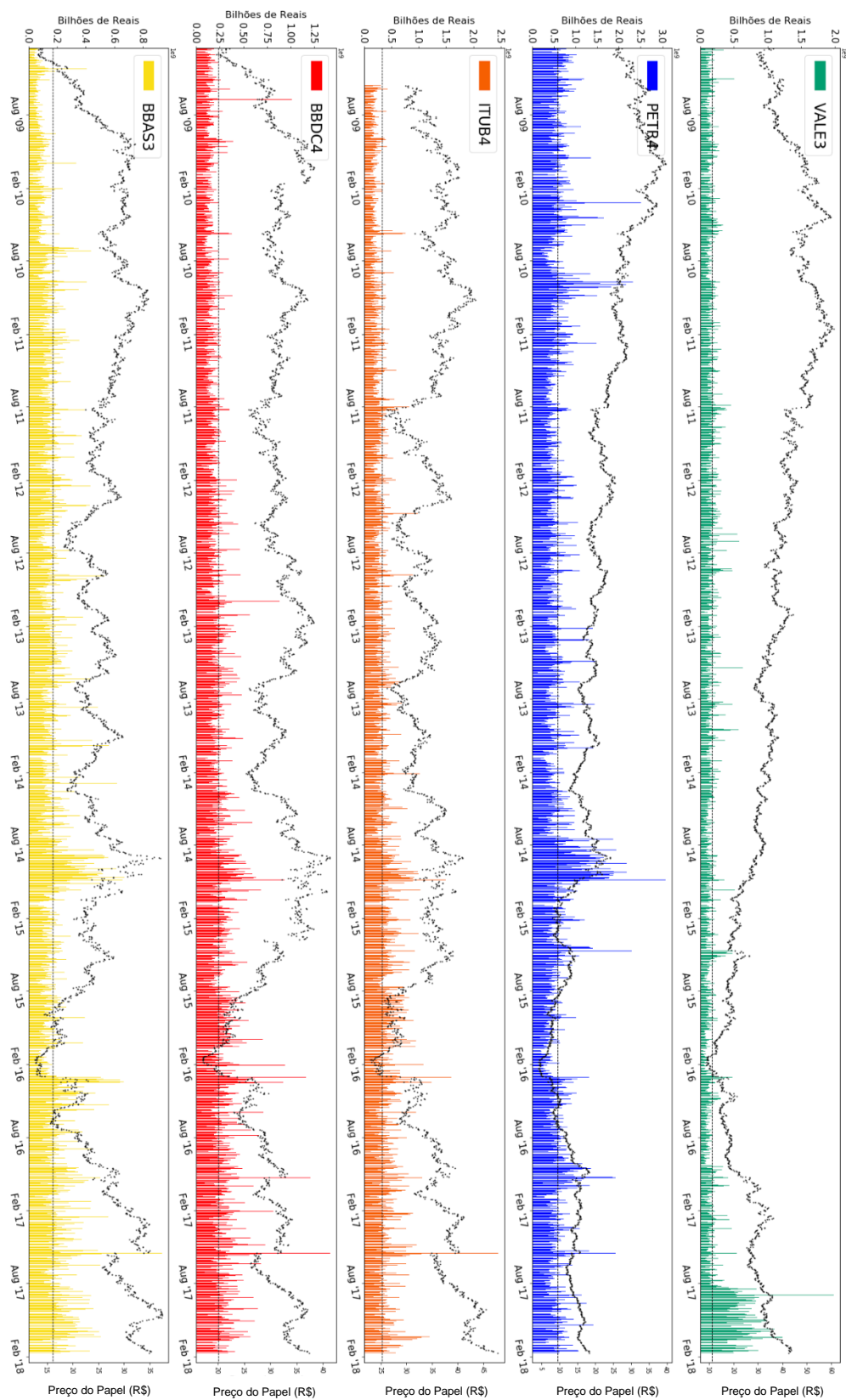
**Fonte:** Elaborado pelo autor

### 3.5 Ibovespa

O Ibovespa, acrônimo para Índice Bovespa, foi criado em 1968 como o indicador do desempenho médio das cotações das ações de maior negociabilidade e representatividade do mercado brasileiro negociadas na BM&FBOVESPA S.A, ele é produto de uma carteira teórica de ativos, composto atualmente por 65 ações de empresas (B3, 2018). Leite e Sanvicente (1994) afirmam que a negociabilidade é o principal fator para a utilização do Ibovespa como índice de mercado, devendo ser considerado para a análise da rentabilidade a variação percentual entre dois pontos históricos quaisquer. A figura 3 mostra a relação de volume financeiro e a evolução do preço das ações que tem maior participação na composição do Ibovespa.



**Figura 3 – Relação de volume financeiro e evolução do preço das ações VALE3, PETR4, ITUB4, BBDC4 e BBSA3**



**Fonte:** Elaborado pelo autor.

### 3.6 Análise fundamentalista

O termo análise fundamentalista é definido como o estudo de valor intrínseco de uma determinada ação e da evolução da empresa associada a fim de determinar se uma empresa é digna de investimento (Schwager e Turner, 1995). Este tipo de análise é associada a demonstrações financeiras (ativos, passivos, balanço, lucro e saúde) de uma empresa. Outros fatores também podem ser considerados como: O estado geral da economia e os fatores, incluindo taxas de juros, produção, ganhos, emprego, PIB, habitação e a situação de mercados específicos relevantes ao negócio da empresa em análise. Kerr (2011), define que a análise fundamentalista consiste em estudar a situação financeira da empresa, considerando os elementos internos, como balanço patrimonial, demonstração do resultado de exercício, fluxo de caixa, entre outros; E externos à empresa, como dados sobre o setor econômico de atuação e sobre o mercado acionário.

O termo fundamentalista é associado a análise de demonstrativos financeiros e outros atributos internos da empresa que são usados para estimar um preço da ação sem interferência do mercado para comparação. Caso esse valor esteja mais baixo que o valor de mercado das ações que é comercializado na bolsa, indica-se a venda do papel, caso esteja acima do valor de mercado, indica-se a compra do papel para valorização e posterior venda. A decisão da compra de uma ação pelo investidor que usa esta técnica implica na análise do potencial lucro associado ao momento da empresa.

Segundo Bauman (1996, apud Wafi, Hassan e Mabrouk, 2015) o verdadeiro começo das pesquisas que atribuem o preço das ações usando análise fundamentalista vem de três artigos principais. O primeiro estudo de Lev e Ohlson (1982) que identificou a necessidade de criar modelos de avaliação de ações para complementar e expandir os estudos de correlação tradicionais para pesquisa contábil com base no mercado. O Segundo estudo foi um artigo publicado por Lev (1989) onde a conclusão é a de que a nossa compreensão do processo de financiamento análise de declaração "não excede muito da lista de índices financeiros, que é suposto ser usado por investidores", e discutiu a necessidade de mudar a pesquisa de mercado de capitais. Por último o artigo de Bernard (1994) que criticou a pesquisa empírica sobre o papel dos dados contábeis na avaliação de ações, entre as sugestões para

pesquisas futuras destaca-se o pedido de estudos que usam amostras de empresas particulares de setores industriais ou econômicos.

De acordo com Moore (2012), alguns exemplos de formas para a projeção de preços da análise fundamentalista são:

1. P/L (Preço/Lucro): preço da ação no mercado / lucro por ação.
2. RPL (Retorno sobre patrimônio líquido): lucro líquido da empresa / patrimônio líquido.
3. LPA (Lucro por ação): lucro líquido / quantidade de ações emitidas.
4. P/VPA (Preço da ação / valor patrimonial por ação).
5. Dividend(%): dividendos / preço da ação.
6. Ebitda: Earning Before Interest, Taxes, Depreciation and Amortization.
7. VE/Ebitda (Valor de empresa sobre Ebitda).

Chen, Chen e Lu (2017) consideram em seu estudo sobre o mercado de ações tailandês informações sobre condições financeiras das empresas, a situação do ambiente industrial correspondente, fatores macroeconômicos e notícias para desenvolver uma abordagem baseada na análise fundamentalista para a previsão da direção das ações.

As projeções de preços calculados na análise fundamentalista, de maneira geral, não refletem as oscilações diárias e são baseados em fatores internos das empresas avaliadas. Neste tipo de análise não é possível verificar todas as variações que o preço de uma ação pode ter durante o seu período de análise.

### 3.7 Análise técnica

Complementar ao método fundamentalista existe a análise técnica que segundo Nazário et al.(2017) pode ser entendida como um conjunto de regras ou gráficos que tende a antecipar o preço futuro, as mudanças são baseadas no estudo de certas informações, como por exemplo: Preço de compra e preço de volume negociado. Outras informações como dados históricos de ações, dados técnicos (indicadores, osciladores e princípios específicos) são usados pelos investidores para prever tendências (Murphy, 1999; Kirkpatrick e Dahlquist, 2010).

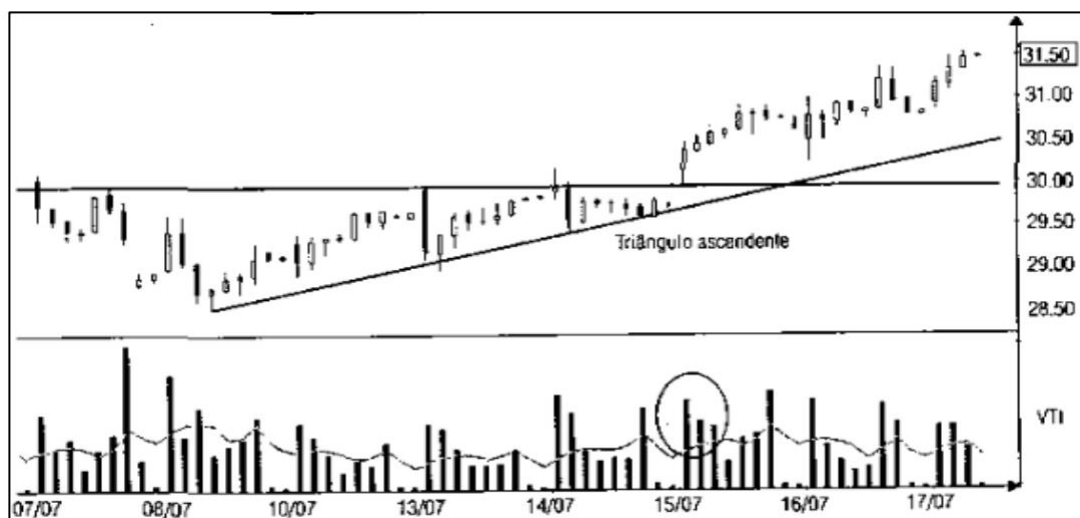
Tendências neste contexto refere-se a representação da direção do preço de uma ação associada ao comportamento dos investidores (mercado). A análise técnica possibilita a identificação de tendências e padrões de comportamento e sua visualização.

Segundo Murphy (2000), enquanto uma linha de tendência não é rompida ela pode ser usada para determinar áreas de compra e venda. Todavia, ele ressalta que o rompimento de uma linha é um dos melhores sinais indicando uma mudança de tendência.

A significância de uma linha de tendência é determinada por dois parâmetros: por quanto tempo ela se manteve intacta e quantas vezes ela foi testada. Uma linha que foi testada oito vezes, por exemplo, e que continuamente demonstrou sua validade, tem muito mais significância do que uma linha que foi tocada apenas três vezes. Adicionalmente, uma tendência que esteja em progresso por nove meses tem muito mais significado do que outra que assim esteja por apenas poucas semanas (Murphy, 1986).

As linhas de tendência não têm fator preditivo, mas ajudam a demonstrar a oscilação do mercado, por exemplo, conforme figura 4 é possível verificar a tendência de valorização da ação PETR4 através do padrão gráfico conhecido como triângulo ascendente que representa um momento de consolidação do preço ao longo do tempo. Observando o quadro abaixo em uma situação de avaliação preventiva, o investidor poderia concluir que o lucro era possível, no entanto este cenário pode não se efetivar. Neste contexto a análise técnica não prevê uma situação e sim a demonstração de uma tendência.

**Figura 4** - Exemplo de linhas de tendência da ação PETR4



**Fonte:** Lemos e Cardoso (2010)

Wang e Chang (2007), Royo, Guijarro e Michniuk (2015) e Arévalo et al. (2017) exemplificam a identificação de padrões na análise técnica do mercado de ações. Os resultados destas pesquisas fornecem evidências que confrontam a hipótese do mercado eficiente de Fama (1964), neste contexto o mercado de ações é previsível e possibilita a formulação de hipóteses para melhor momento de compra e venda de ações.

Namdari e Li (2017) vão além, e em seu experimento e combinam a análise técnica e fundamentalista na previsão de ações através de índices financeiros das empresas e dados históricos de ações para prever o movimento das ações, para isto foi utilizado redes neurais artificiais do tipo multilayer perceptron. O estudo de Namdari e Li (2017) valida duas hipóteses: Primeiro o uso de redes neurais para a previsão da direção das ações comercializadas no mercado financeiro. Segundo, a conclusão apresentada por Bettman et al.(2009) de que as análises técnica e fundamentalista são complementares e não excludentes.

### 3.8 Séries temporais

Souza (1989) define uma série temporal como uma classe de fenômenos cujo processo observacional e consequente quantificação numérica geram uma sequência de observações distribuídas ao longo do tempo. Bayer e Souza (2010) estabelecem

uma série temporal como um conjunto de observações ordenadas no tempo de qualquer fenômeno aleatório. A análise de séries temporais consiste em encontrar relações de dependência existentes temporalmente nos dados buscando-se identificar o mecanismo gerador da série, com o objetivo de extrair periodicidades relevantes nas observações, descrever seu comportamento e fazer previsões.

Morenttin e Toloi (2006) descrevem os possíveis objetivos de um estudo associado a uma série temporal:

- Avaliação da natureza e a estrutura de pesquisas anteriores sobre o uso de redes neurais no mercado de ações, para revalidar, contestar ou melhorar métodos e resultados.
- Realizar a previsão de um ou mais passos futuros.
- Descrever o comportamento da série.
- Procurar a periodicidade relevante nos dados.

Segundo Lopes (2013) um processo estocástico é uma família de variáveis aleatórias  $\{X(t), t \in T\}$  definidas em um espaço de probabilidade, indexado por um parâmetro  $t$ , onde  $t$  varia no conjunto  $T$ .

Levine, Berenson e Stephan (2000) afirmam que os elementos do modelo de séries temporais, quando aplicados ao mercado financeiro, podem ser definidos da seguinte maneira:

- A tendência de uma série é sua direção ascendente ou descendente de longo prazo, que indica evolução ou queda ao longo do tempo.
- As variações cíclicas mostram as oscilações ao longo da série. São constituídas por extensões, intensidades e amplitudes variadas. Costumam durar até 10 anos e, geralmente, estão ligadas a um ciclo de negócios.
- As variações sazonais são um caso particular de variações cíclicas, no qual os dados são registrados em prazos curtos, diários, semanais ou mensais.

Conforme Morettin e Tolói (2006), as séries financeiras podem ser classificadas como sendo não-estacionárias. Uma série temporal é considerada estacionária quando apresenta um equilíbrio estável, ou seja, os dados se apresentam ao redor de uma média e com variância constante. Em geral, entretanto, as séries financeiras apresentam tendências. Neste caso, a série se desenvolve ao redor de uma reta com

inclinação positiva ou negativa. Outras características das séries financeiras são: a sazonalidade, os pontos atípicos, a heterocedasticidade condicional e a não-linearidade. Também segundo Morettin e Tolói (2006) a análise de séries financeiras pode ser realizada através de modelos paramétricos e não paramétricos, e os modelos mais adequados para séries financeiras são os modelos não-lineares, característica apresentada pelos modelos de redes neurais artificiais.

### 3.9 Redes neurais artificiais

Muitos autores declaram diferentes conceitos sobre redes neurais artificiais, o que demonstra a profundidade que o assunto foi explorado pela comunidade científica. Azoff (1994) define redes neurais como uma técnica de processamento de dados que mapeia ou relaciona algum tipo de fluxo de entrada de informações para um fluxo de saída de dados. Gurney (1997) determina uma rede neural artificial ou RNA, como um conjunto interconectado de elementos, unidades ou nós de processamentos únicos cuja funcionalidade é baseada em um neurônio animal. A capacidade de processamento da rede é armazenada nos pontos fortes da conexão entre unidades, ou pesos obtidos por um processo de adaptação ou aprendizado de um conjunto de padrões de treinamento. Selli e Seleglim (2007) descrevem em uma análise matemática que as RNAs podem ser definidas como um mapeamento não-linear entre um vetor de espaço de entrada e um vetor de espaço de saída, este mapeamento pode ser feito através de camadas de funções de ativação ou neurônios onde as coordenadas de entrada são somadas de acordo com o valor de seus respectivos pesos para produzir uma saída de acordo com o respectivo nível de estímulo.

Segundo Haykin (2001) uma rede neural artificial, é um processador maciço, paralelamente distribuído e constituído de unidades de processamento simples, que têm a propensão natural de armazenar conhecimento experimental e torná-lo disponível para uso, a semelhança com o cérebro humano é nos seguintes aspectos:

1. O processo de aprendizado para formação do conhecimento acontece a partir da interação com o ambiente.
2. O conhecimento aprendido é armazenado na força de conexão entre os neurônios, também conhecido como peso sináptico.

Conclui-se então que uma rede neural artificial como um algoritmo baseado na capacidade de aprendizado do cérebro humano e no comportamento de suas células específicas, os neurônios. O ser humano é capaz de realizar tarefas consideradas complexas quando comparado a outros animais, por exemplo: Reconhecimento de padrões, classificação e generalização.

De acordo com Trippi and Turban (1996), Schalkof (1997) e Goonatilake e Treleaven (1995) e também Haykin (2001) o uso de redes neurais oferece as seguintes propriedades e capacidades úteis:

1. Não-linearidade. Um neurônio artificial pode ser linear ou não linear. Uma rede neural, feita da interconexão de neurônios não lineares é por si não linear. Além disso, a não-linearidade é de um tipo especial no sentido de que é distribuída por toda a rede. A não-linearidade da rede é uma propriedade importante particularmente se o mecanismo que provê os sinais de entrada também é inerentemente não linear (ex: sinais de fala).
2. Mapeamento entrada-saída. Um dos paradigmas de aprendizado chamado de aprendizado supervisionado que envolve a modificação dos pesos sinápticos da rede neural aplicando um conjunto de amostras de treinamento rotuladas ou exemplos de tarefas. Cada exemplo consiste em um sinal de entrada exclusivo e a resposta desejada correspondente. A rede é apresentada a um exemplo aleatório do conjunto e os pesos sinápticos são modificados para minimizar a diferença entre a resposta desejada e a atual resposta produzida pelo sinal de entrada de acordo com os critérios estatísticos apropriados. O processo de treinamento é repetido com vários exemplos até que a rede atinja um estado em que não há mudanças significativas nos pesos sinápticos. A rede neural artificial aprende com os exemplos através da construção de um mapeamento de entrada-saída.
3. Adaptabilidade. Redes neurais tem uma capacidade interna de adaptar seus pesos sinápticos as mudanças do ambiente. Particularmente uma rede neural treinada para operar em um ambiente específico pode facilmente ser retreinada para lidar com pequenas alterações nas condições ambientais operacionais.



4. Resposta evidencial. No contexto de classificação de padrões, uma rede neural pode ser projetada para prover informação não somente sobre o padrão particular selecionado, mas também sobre a confiança na decisão tomada
5. Informação contextual. O conhecimento é representado pela própria estrutura e estado de ativação de uma rede neural. Cada neurônio da rede potencialmente é afetado pela atividade de todos os demais neurônios da rede. Consequentemente a informação contextual é tratada naturalmente pela rede neural.
6. Tolerante a falhas. Uma rede neural implementada em hardware tem o potencial para ser tolerante a falhas e computacionalmente capaz, no sentido de que seu desempenho diminua graciosamente sob condições adversas de operação.
7. Uniformidade de análise e design. A mesma notação é usada em todos os domínios que envolvem redes neurais, esse recurso pode ser observado de diferentes maneiras:
  - a. Neurônios são presentes em todas as redes neurais.
  - b. Este fator comum possibilita o compartilhamento de teorias e algoritmos de aprendizado em diferentes aplicações de redes neurais.
  - c. Redes modulares podem ser construídas através de uma integração perfeita de módulos.
8. Analogia neurobiológica. O design de uma rede neural artificial é motivado pela analogia com o cérebro humano, prova de um processamento paralelo poderoso e rápido.

Relativo ao presente estudo foram identificadas pesquisas que aplicaram redes neurais para previsão de ações, White (1988), Lam (2004), Oliveira, Nobre e Zárte (2013), Silva et al. (2014), Altinbas e Biskin (2015), Zahedi e Rounaghi (2015), Mingyue, Cheng e Yu (2016), Oriani e Coelho (2016), Gunduz, Yaslan e Cataltepe (2017), Nelson, Pereira e Oliveira (2017), Samarawickrama e Fernando (2017) e Xiong e Lu (2017). Estes estudos concluem que redes neurais artificiais podem ser utilizadas para previsão de preço e da direção de ações no mercado financeiro.

### 3.9.1 Contexto histórico

O surgimento das redes neurais artificiais deu-se com o modelo matemático do neurônio biológico proposto por Warren McCulloch e Walter Pitts em 1943, o modelo proposto é baseado na unidade mais simples de um sistema neural, o neurônio. Onde **N** entradas são multiplicadas por um peso **X** e seus resultados são comparados a um limiar.

Hebb (1949), demonstrou o papel de circuitos reverbatórios no cérebro e como eles poderiam ser usados como memória de curta duração. Relacionado ao modelo de neurônio de McCulloch e Pitts, Hebb mostrou como a aprendizagem de uma rede neural é variante através dos pesos de entrada nos neurônios, processo que hoje é conhecido como aprendizagem Hebbiana. Segundo Hebb (1949) tem-se:

$$\Delta w_{ij} = nx_i o_j \quad (1)$$

Onde:

$w_{ij}$  é a intensidade da conexão entre os neurônios  $i$  e  $j$

$\Delta w_{ij}$  é o acréscimo da intensidade da conexão entre os neurônios  $i$  e  $j$

$n$  parâmetro que define a intensidade da correção, taxa de aprendizado

$x_i$  estado de ativação do neurônio  $i$

$o_j$  saída do neurônio  $j$

Rosemblatt (1958) propôs uma topologia denominada Perceptron, com duas camadas e neurônios do tipo McCulloch e Pitts. No modelo proposto não existia uma camada oculta ou intermediária, o processo de aprendizado acontecia na mesma camada de saída. Este estudo representa os primeiros esboços de uma rede neural simulada com mais de uma camada. Alguns anos depois Widrow (1962) propôs um modelo chamado Adaline, acrônimo para "Adaptative Linear Element". Neste modelo, o ajuste dos pesos da rede é realizado através de uma aproximação linear do cálculo do gradiente de uma função de erro quadrática, conhecida como Regra de Widrow que pode ser expressa conforme abaixo:

$$E(t) = (y^d - y(x)) \quad (2)$$

$$W(t+1) = W(t) - \eta \frac{\partial E}{\partial t} \quad (3)$$

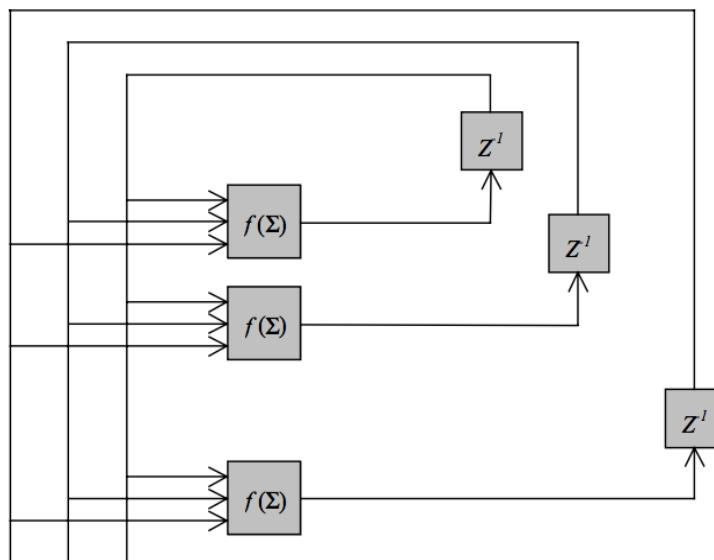
Onde  $t$  é a iteração do treinamento,  $\eta$  é o fator de aprendizagem,  $y^d$  é a saída desejada,  $y(x)$  é a saída do neurônio,  $W(t)$  são os pesos na iteração  $t$  e  $\frac{\partial E(t)}{\partial t}$  é o gradiente do erro na iteração  $t$ .

Minsky e Papert (1969) publicaram um livro chamado "Perceptrons". Neste livro as capacidades adaptativas e computacionais dos modelos de neurônios do tipo perceptron foram analisadas e concluiu-se que apesar do modelo ser capaz de executar operações booleanas AND e OR ele não era capaz de executar outros tipos de operações também consideradas elementares, por exemplo o XOR. Os autores também não acreditavam que esta limitação pudesse ser superada. Após a publicação destes resultados as pesquisas em redes neurais tiveram uma perda significativa de financiamentos e de pesquisadores. Outro importante fator para esta queda de trabalhos em redes neurais foi a grande expectativa criada pelos pesquisadores da área, não acompanhada de resultados à altura.

Na década de 80, alguns fatores contribuíram para um foco maior nas pesquisas sobre redes neurais quando comparados à década anterior:

1. Muitos avanços tecnológicos na área da computação possibilitaram maior poder computacional a baixo custo, o que tornou possível simulações mais sofisticadas e que pudessem replicadas com mais facilidades por mais pesquisadores.
2. O desenvolvimento de novas teorias para a implementação de algoritmos adaptativos.

A pesquisa de Tank e Hopfield (1982) pode ser explicada como uma rede dinâmica com um número finito de estados que podem ser utilizados como dispositivos de memória. O acesso ao conteúdo da memória de uma rede de Hopfield se dá permitindo que a rede evolua com o tempo para um de seus estados de memória, denominados memórias endereçáveis por conteúdo.

**Figura 5** - Representação de uma rede Hopfield

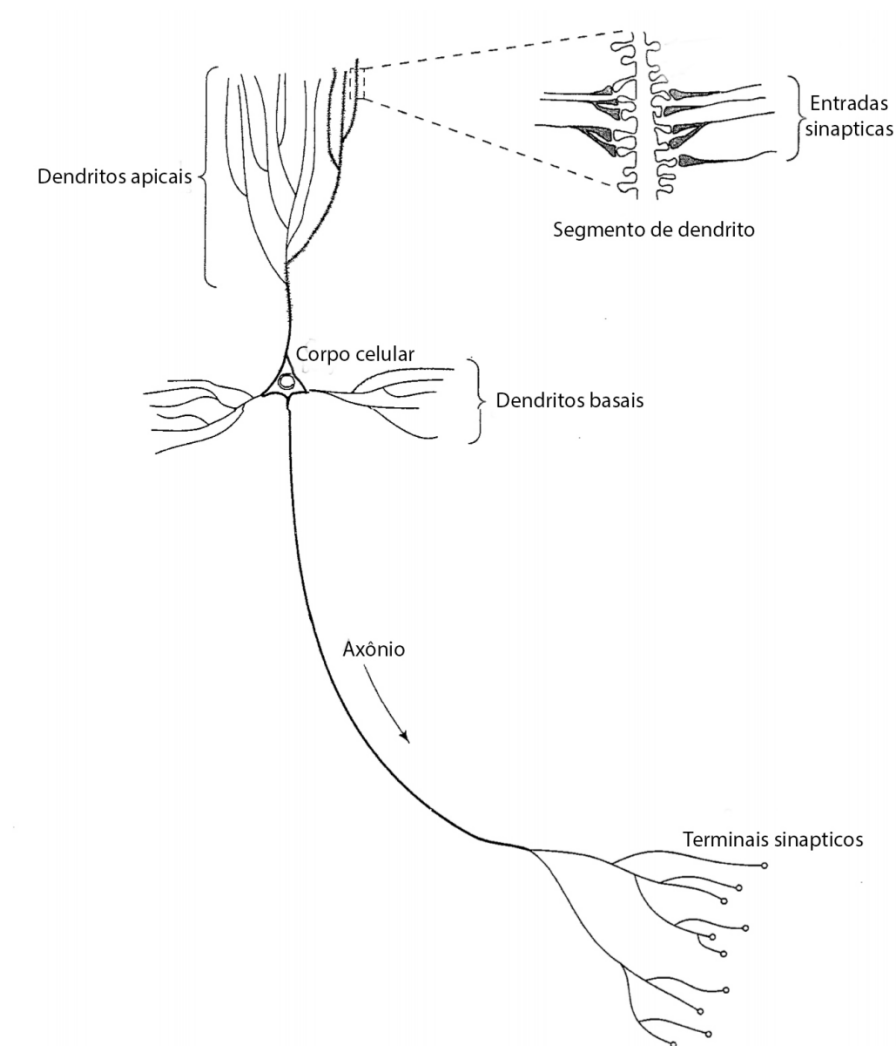
**Fonte:** Elaborado pelo autor

Rumelhart et al. (1986), publicaram um método para ajuste de parâmetros de redes não-recorrentes de múltiplas camadas conhecido hoje como retropropagação. Este algoritmo possibilitou a evolução do modelo proposto por Roseblatt (1958) com redes mais complexas aumentando o poder de generalização das redes neurais. Isso também propiciou o treinamento de redes Perceptron com mais de uma camada, este estudo tornou-se uma referência para as pesquisas sobre redes neurais, possibilitando a implantação de diversas possibilidades.

### 3.9.2 Estrutura do neurônio

A figura 6 demonstra um neurônio biológico que se conecta com outros neurônios através de sinapses. O tipo mais comum de sinapse é a sinapse química, que opera da seguinte forma: Um processo pré-sináptico libera uma substância transmissora que se difunde através da junção sináptica entre neurônios e então age sobre um processo pós-sináptico. Assim, uma sinapse converte um sinal elétrico pré-sináptico em um sinal químico e então de volta em um sinal elétrico pós-sináptico (Sheperd e Koch, 1990). Vale ressaltar que essa conexão pode impor ao neurônio receptivo excitação ou inibição, mas não ambas.

**Figura 6** - Representação de um neurônio biológico



**Fonte:** Haykin (2001)

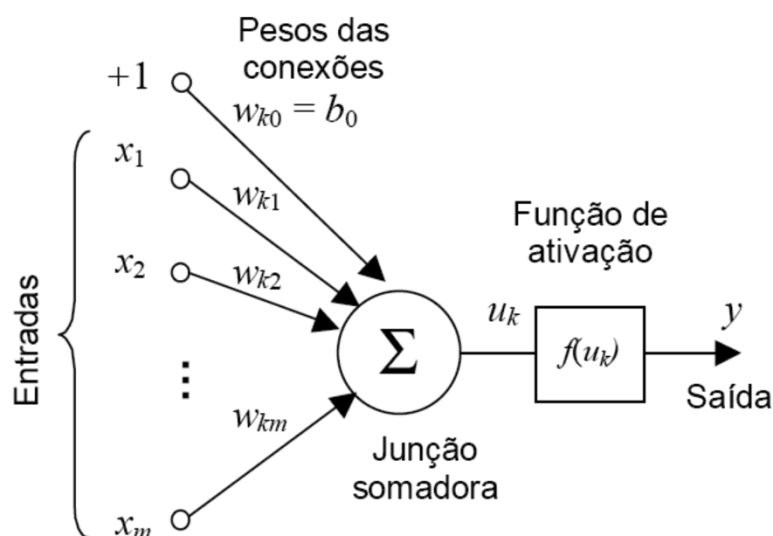
Os dendritos são prolongamentos ramificados com um formato parecido ao de uma árvore. Sua função é receber os estímulos de vários outros neurônios.

O corpo celular é onde ficam as principais organelas celulares do neurônio, esta parte é responsável pelo processamento de toda a informação que chega dos dendritos, pela produção de um potencial de ativação que indica se o neurônio pode produzir um impulso elétrico junto com seu axônio.

Por fim, o axônio, é um único prolongamento cuja principal função é guiar os impulsos elétricos para outros neurônios de conexão. A terminação do axônio é também composta de ramos chamados terminais sinápticos.

Como pode ser observado na figura 7, a menor unidade de uma rede neural artificial é o neurônio. Estas unidades, são dispositivos não-lineares e adaptáveis, embora muito simples em termos de poder computacional e de memória. Porém, quando interligadas, apresentam um enorme potencial para mapeamentos não-lineares, similar ao cérebro humano.

**Figura 7** - Representação matemática de um neurônio do tipo Perceptron



**Fonte:** Elaborado pelo autor

Onde  $x_1$ ,  $x_2$  e  $x_m$  são as representações dos sinais de entrada ou dados do algoritmo.  $w_{k0}$ ,  $w_{k1}$ ,  $w_{k2}$  e  $w_{km}$  são os pesos sinápticos que são aprendidos durante o processo de treinamento, o viés representado por  $b_0$  também está relacionado aos pesos. A junção somadora ou combinador linear representado por  $\Sigma$  serve como agregador de todos sinais de entrada que foram ponderados pelos respectivos pesos sinápticos a fim de produzir um potencial de ativação representado por  $u_k$  que é o resultado obtido pela diferença do valor produzido entre a junção somadora  $\Sigma$  e um limiar de ativação para uma função de ativação  $f(u_k)$  que limita o resultado do processamento de um neurônio a um intervalo de valores a uma saída  $y$  que pode estar ligada a um outro neurônio em uma rede, por exemplo.

Cada unidade de processamento, que simula um neurônio, possui ligações com outras unidades que recebem e transmitem sinais. Uma rede neural pode possuir uma ou múltiplas camadas de neurônios artificiais que se interconectam. O algoritmo

de aprendizagem é o procedimento utilizado para realizar o processo de aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado (Haykin, 2001).

### 3.9.3 Características de redes neurais artificiais

Com a evolução do poder de processamento, o conceito simples de um perceptron pôde ser explorado e evoluído para uma rede neural artificial ou RNA (acrônimo de rede neural artificial). Atualmente redes neurais artificiais são utilizadas em diversas áreas do conhecimento humano e possuem características que são baseadas nas células nervosas do ser-humano. Sobre redes neurais artificiais é possível assumir que:

1. A propagação dos sinais de um elemento a outro que representa um neurônio acontece através de conexões baseadas nos dendritos e nos axônios.
2. O processamento dos dados acontece nos elementos que representam o corpo celular do neurônio.
3. Cada conexão possui um peso associado, que, em uma rede neural típica, pondera o sinal transmitido.
4. Cada elemento aplica uma função de ativação à sua entrada de rede (soma ponderada dos sinais de entrada) para determinar sua saída.

Uma rede neural pode ser caracterizada por três aspectos principais: Arquitetura da rede, o algoritmo de treinamento ou aprendizagem e a função de ativação.

### 3.9.4 Função de ativação

A Função de ativação que segundo Haykin (2001) pode ser definida como a representação do efeito que a entrada e o estado atual da ativação representam na ativação da próxima unidade. Ela representa o efeito que a entrada interna e o estado atual de ativação exercem na definição do próximo estado de ativação da unidade. Quando propriedades dinâmicas estão envolvidas na definição do estado de ativação, equações diferenciais (caso contínuo) ou as diferenças (caso discreto) são empregadas. Geralmente define-se o estado de ativação de um neurônio através de

uma função algébrica e apresenta algum tipo de não-linearidade. Alguns dos tipos mais utilizados de função de ativação são:

- Função linear: Função linear que simplesmente repete o sinal que entra no neurônio na sua saída:

$$f(x) = p.x, \quad f'(x) = p \quad (4)$$

- Função degrau: Modelo de ativação tudo-ou-nada, pode ser expresso pela equação abaixo:

$$f(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases} \quad (5)$$

- Função sigmóide: A função sigmóide ou logística possui intervalo de variação entre 0 e 1, é uma função contínua. A origem desta função está vinculada à preocupação em limitar o intervalo de variação da derivada da função, pela inclusão de um efeito de saturação. Sua derivada também é uma função contínua:

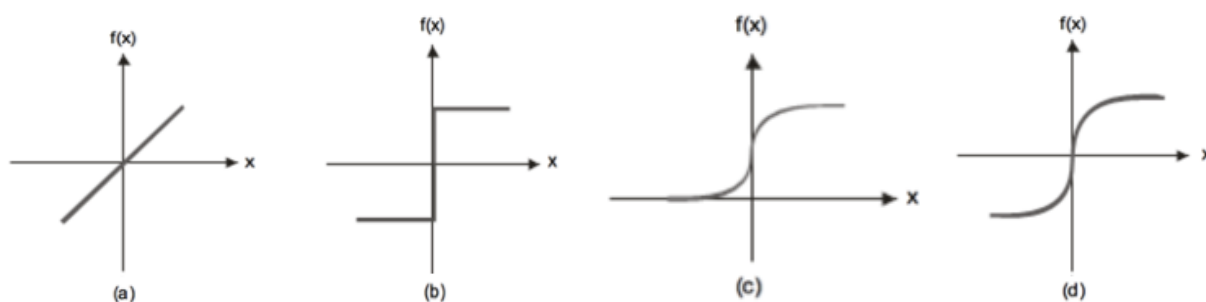
$$f(x) = \frac{e^{px}}{1 + e^{px}} = \frac{1}{1 + e^{-px}}, \quad f'(x) = p f(x) \cdot (1 - f(x)) \quad (6)$$

- Função tangente hiperbólica: Pode ser considerada como uma evolução da função sigmóide, a função tangente hiperbólica apresenta a mesma forma sigmoide e também valores positivos e negativos. A função de tangente hiperbólica e sua respectiva derivada podem ser expressas como:

$$f(x) = \frac{e^{px} - e^{-px}}{e^{px} + e^{-px}} = \tanh(px), \quad f'(x) = p(1 - f(x)^2) \quad (7)$$



**Figura 8** - Funções de ativação: Linear, degrau, sigmóide e tangente hiperbólica



**Fonte:** Elaborado pelo autor

### 3.9.5 Arquitetura de redes neurais

A arquitetura ou topologia das redes neurais refere-se a quantidade de camadas, neurônios e como eles se conectam para que o fluxo de informações (entrada e saída) ajuste e altere o comportamento da rede. As unidades de uma rede neural são conectadas por canais de comunicação associadas a um peso. Se o aprendizado está nos pesos das conexões, o principal objetivo é identificar um conjunto apropriado de pesos de forma que a rede se comporte como desejado.

Os fatores devem ser especificados de acordo com o problema a ser analisado. A elaboração de uma rede neural envolve: A quantidade de camadas que a rede deverá ter (entrada, intermediárias e de saída), o número de neurônios da rede por camada, os dados de entrada, a apresentação da saída e a função de ativação dos neurônios. O número de nós na camada de entrada é determinado pela complexidade do problema a ser resolvido. A camada de saída é determinada pela resposta desejada.

Dentre as arquiteturas de redes neurais artificiais destacam-se as redes neurais do tipo multilayer perceptron Rumelhart et al. (1986), as redes neurais convolucionais (Le Cun et al., 1998) e as redes neurais recorrentes Tank e Hopfield (1982). Os estudos de Hopfield (1982) foram os precursores de outro tipo de arquitetura de rede neural também importante a Long Short-Term Memory (Hochreiter e Schmidhuber, 1997).

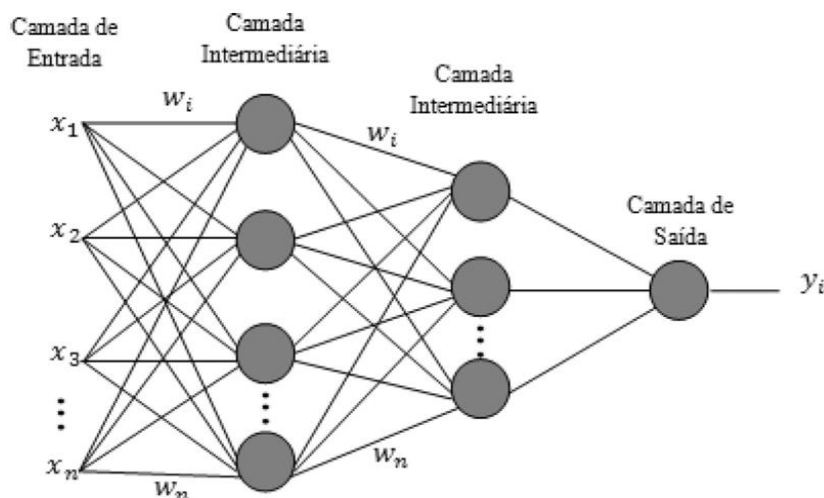
#### 3.9.5.1 Rede multilayer perceptron

A rede multilayer perceptron ou MLP é uma generalização do perceptron apresentado anteriormente, tipicamente uma rede neural deste tipo tem como entrada

o resultado de processamento ou saída da camada anterior, caracterizando um processo feedforward ou acíclico, o fluxo dos dados segue apenas uma única direção. A rede apresentada é totalmente interconectada, o que significa que uma unidade (neurônio) em qualquer camada da rede está conectada a todas as outras unidades (neurônios) na camada anterior. Segundo Haykin (2001). Este tipo de rede tem as seguintes características:

1. Camada de entrada, onde os as variáveis e os padrões são apresentados ao algoritmo.
2. Camada(s) intermediária ou camada oculta, onde o aprendizado e a identificação de padrões acontece através das conexões (sinapses) entre os neurônios.
3. Camada de saída ou apresentação, onde o resultado do processamento é apresentado.

**Figura 9** – Representação de rede neural multilayer perceptron



**Fonte:** Elaborado pelo autor

Algumas características das redes neurais do tipo MLP são:

- Este tipo de rede neural é composta por no mínimo uma camada oculta, estas camadas internas possibilitam que a rede possa aprender tarefas mais complexas.
- Cada neurônio possui uma função de ativação do tipo não-linear. A função de ativação mais comum em estudos de redes neurais é do tipo sigmóide ou sigmoidal.

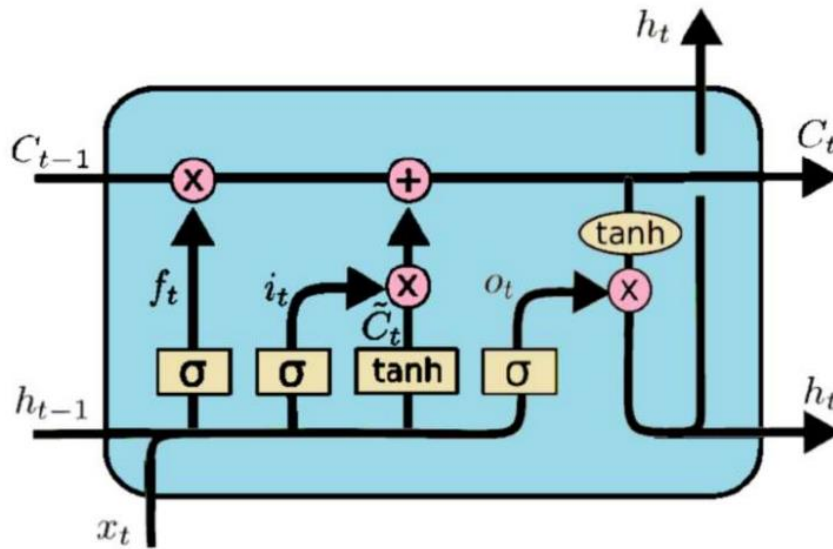
- Uma rede MLP é uma rede do tipo feedforward ou acíclica. Os dados seguem sempre um único fluxo. As saídas dos neurônios de uma camada se conectam unicamente as entradas dos neurônios das camadas seguintes.
- A rede tem um alto grau de conectividade entre seus elementos (neurônios) representando as sinapses. A conectividade é alterada por uma mudança no número de conexões ou no peso delas.

Este trabalho utilizará redes neurais artificiais para prever o preço de fechamento de ações. Lam (2004) sugere o uso de redes neurais artificiais para previsões no mercado financeiro pelos seguintes motivos: Primeiro, redes neurais não requerem nenhuma suposição de distribuição de dados, recurso que torna este tipo de algoritmo mais flexível do que técnicas estatísticas como a análise de regressão, por exemplo. Segundo, ela permite que novos dados sejam submetidos a uma rede neural já treinada para atualizar o resultado do treinamento anterior, sem a necessidade destes novos dados estarem misturados aos já processados, o que a torna ideal para aplicações financeiras com novos dados atualizados rapidamente.

### 3.9.5.2 Redes LSTM

Hochreiter e Schmidhuber (1997) citam que em redes neurais que utilizam gradientes, por exemplo a retropropagação convencional em uma rede do tipo multilayer perceptron através do tempo ou aprendizado recorrente tendem a explodir ou a desaparecer, o que pode tornar a rede desbalanceada no caso da explosão ou incapaz de aprender dependências após muitos estágios de processamento. A topologia LSTM ou Long Short-Term Memory proposta por estes pesquisadores trabalha com estas limitações e a torna capaz de processar dependências longas. A pesquisa de Gref, et al (2015) faz uma comparação entre oito adaptações de LSTM e conclui que o modelo de Hochreiter e Schmidhuber (1997) tem um desempenho aceitável quando comparado a modelos mais recentes.

**Figura 10** – Representação de uma célula de uma rede LSTM



**Fonte:** Adaptado de OLAH, C (2015)

A figura 10 representa a estrutura de uma célula LSTM onde é possível identificar os componentes denominados portões e conexões. O termo  $f_t$  é o portão de esquecimento que após receber o estado externo  $h_{t-1}$  filtra o que não é mais necessário ao estado  $h_{t-1}$  através de uma função sigmoide e pode ser representado conforme abaixo:

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

No próximo passo outra camada sigmoide é responsável pela decisão de quais valores devem ser atualizados e a camada  $\tanh$  (tangente hiperbólica) cria um vetor com novos candidatos ao presente estado da célula.

$$\begin{aligned} i_t &= \sigma (W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \end{aligned} \quad (9)$$

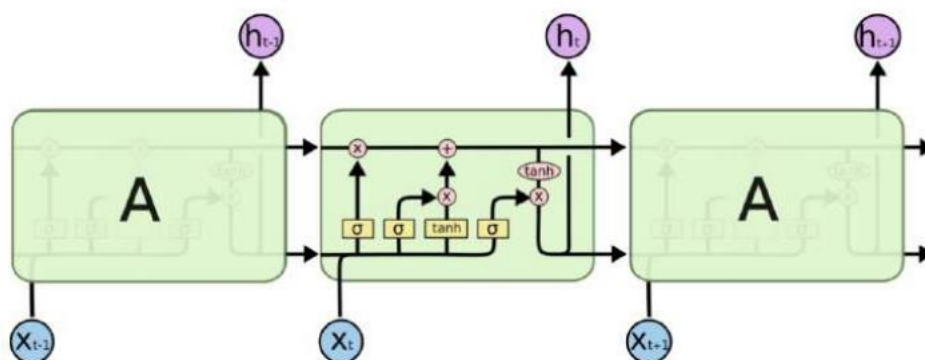
Por fim os resultados do estado da célula são atualizados por outra função sigmoide em conjunto com outra função  $\tanh$ .

$$\begin{aligned} o_t &= \sigma (W_o [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh (C_t) \end{aligned} \quad (10)$$

Também é possível observar os pontos de conexão com o exterior da célula representados por:  $C_{t-1}$ ,  $h_{t-1}$ ,  $x_t$ ,  $C_t$  e  $h_t$ . Estes pontos representam o estado da

célula no instante  $t$  o momento em que as informações chegaram até este instante de tempo em outras iterações. Uma rede neural LSTM geralmente é composta por mais de uma célula conforme exposto na figura 11.

**Figura 11** – Representação de uma rede LSTM composta por 3 células



**Fonte:** Adaptado de OLAH, C (2015)

### 3.10 O processo de aprendizagem de redes neurais artificiais

Mendel e McClaren (1970) definem o processo de aprendizagem para redes neurais como um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem então é determinado pela maneira pela qual a modificação dos parâmetros ocorre.

Segundo Haykin (2001) esta definição implica a sequência de eventos:

1. A rede neural é estimulada por um ambiente.
2. A rede neural sofre modificações nos seus parâmetros livres como resultado desta estimulação.
3. A rede neural responde de uma maneira nova ao ambiente, devido às modificações ocorridas na sua estrutura interna.

Haykin (2001) classifica um conjunto preestabelecido de regras bem-definidas para a solução de um problema de aprendizagem como um algoritmo de aprendizagem. Não há um algoritmo de aprendizagem único para o projeto de redes neurais. Em vez disso há um conjunto de ferramentas representado por uma variedade de algoritmos de aprendizagem, cada qual com vantagens específicas.

O algoritmo de aprendizagem ou treinamento pode ser dividido em 2 tipos, supervisionado e não-supervisionado. A aprendizagem supervisionada necessita de um “professor” ou “guia” durante a fase de aprendizagem, antes do uso da rede neural. Já o não-supervisionado é direcionado por correlações existentes nos dados de entrada e o resultado esperado, portanto, ele “aprende sozinho”. Alguns exemplos de redes neurais artificiais separadas pelo método de treinamento são:

Aprendizado supervisionado:

- Perceptron
- Adaline
- Retropropagação – (Backpropagation)
- Rede lógica adaptativa
- Rede neural probabilística
- Máquina de Boltzmann
- Mean field annealing

Aprendizado não-supervisionado:

- Matriz de aprendizado (Learning matrix)
- Memória associativa linear
- Aprendizado competitivo
- Memória associativa bidirecional
- Memória associativa temporal

### 3.11 O algoritmo de retropropagação e outros algoritmos de treinamento

A aprendizagem de redes MLP por retropropagação consiste de duas etapas através das diferentes camadas da rede: A propagação e a retropropagação. Na etapa de propagação, um padrão de ativação é aplicado aos nós da camada de entrada da rede e seu efeito se propaga através da rede, camada por camada. Na última camada, um conjunto de saídas é produzido como resposta real da rede. Deve-se salientar que, durante a etapa de propagação, os pesos sinápticos da rede são todos fixos. Durante a etapa de retropropagação, os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro (treinamento supervisionado). A resposta real da rede

é subtraída de uma resposta desejada para produzir um sinal de erro. Este sinal de erro é então propagado para trás através da rede, contra a direção das conexões, sendo os pesos sinápticos ajustados para fazer com que a resposta real da rede se aproxime da resposta desejada.

O uso do algoritmo de retropropagação implica na determinação de um ganho ou taxa de aprendizado para o treinamento da rede. A escolha de um ganho fixo para este treinamento deve ser feita com cuidado. Se este ganho for grande, o aprendizado pode ocorrer rapidamente, mas ele pode também se tornar instável. De outro modo, se o ganho for suficientemente pequeno, pode-se garantir estabilidade no treinamento, mas existe o risco de haver um tempo de treinamento muito longo. A escolha deste ganho pode se tornar ainda mais difícil se considerar inicializações diferentes para os pesos da rede e diferentes topologias de redes MLP.

Posteriormente, foram utilizadas técnicas que aperfeiçoaram o algoritmo de retropropagação na otimização do ajuste dos pesos. A revisão sistemática de Tkáč e Verner (2016) apresenta alguns destes algoritmos de treinamento de redes neurais mais utilizados: BA: Bayesian learning (aprendizado bayesiano), CL: competitive learning (aprendizado competitivo), CG: conjugate gradient (conjugado gradiente), GD: gradient descent (gradiente descendente), HE: Hebbian learning (aprendizagem hebbiana), HY: hybrid learning (método de aprendizado híbrido), LM: Levenberg-Marquardt algorithm (método de Levenberg–Marquardt) (Marquardt, 1963), QN: Quasi-Newton algorithms (métodos Quasi-Newton).

Tkáč e Verner (2016) evidenciam a atenção da comunidade acadêmica sobre a aplicação de redes neurais no mercado financeiro nas seguintes áreas: Auditoria e contabilidade, monitoramento de custos, escore de crédito, métricas de consumidores, suporte a decisão, derivativos, câmbio e taxa de juros, análises financeiras, dificuldades financeiras e falência, análise de fraude, inflação, marketing, outras revisões sistemáticas, vendas, títulos e ações. No período 1994 a 2015 o artigo avaliou de forma empírica 412 artigos.

A revisão sistemática de Tkáč e Verner (2016) aponta que a topologia mais utilizada para o desenvolvimento das pesquisas foi a multilayer feedforward e o algoritmo de treinamento mais utilizado foi o gradiente descendente. Os autores indicam que os esforços futuros devem focar em diretrizes universais e metodologia geral para o estabelecimento de variáveis de controle, seleção de camadas ocultas e

desenho geral da topologia, há também o destaque na cautela da análise dos resultados pelos pesquisadores que usam a validação das técnicas comparadas a métodos convencionais.

### 3.12 Trabalhos relacionados

O Quadro 1 é apresenta um resumo comparativo dos trabalhos considerados válidos para o estudo desenvolvido, com ênfase nos métodos ou técnicas usadas nas pesquisas além dos principais conceitos empregados.

**Quadro 1 - Trabalhos relacionados**

Trabalho relacionado	As ações de mais de uma empresa foram avaliadas?	Variáveis Macroeconômicas foram usadas?	Topologia da rede neural usada	Objetivo da pesquisa
Sousa (2016)	Sim	Sim	N/A	Análise de fatores macroeconômicos na previsão de retorno de ativos nos BRICS
Oliveira, Nobre e Zárate (2013)	Não, somente a ação PETR4.	Sim	Multilayer Perceptron	Estudo de caso da ação PETR4, para verificar o uso de redes neurais para previsão do preço.
Gunduz, Yaslan e Cataltepe (2017)	Sim, os dados de 100 ações comercializadas na BIST (Bolsa de valores de Istambul) foram utilizados neste experimento.	Não	Rede Neural Convolutacional	Prever a direção horária de 100 ações Borsa Istanbul Stock Market.
Zahedi e Rounaghi (2015)	Sim, a pesquisa incluiu as empresas listadas na Bolsa de Teerã no período de 2006 até o final de março de 2012.	Não	Multilayer Perceptron	Influência de variáveis contábeis no uso de redes neurais para a previsão do preço da bolsa de Teerã.
Silva, Castilho, Pereira (2014)	Sim, as ações BBDC4, BBSA3, USIM5, PETR4 e VALE5	Não	Multilayer Perceptron	Prever a oscilação dos preços das ações.
Altinbas e Biskin (2015)	Não, apenas as ações correspondentes ao BIST100.	Sim	Multilayer Perceptron	Identificar os indicadores macroeconômicos que tem influência nos movimentos do mercado turco
Samarawickrama e Fernando (2017)	Sim, ações das empresas Commercial Bank Plc. (COM), Royal Ceramics Limited (RCL) e John Keels Holdings (JKH)	Não	Híbrida: SRNN, (LSTM) e GRU	Estudar a aplicabilidade de redes neurais recorrentes para a previsão de preços de ações na bolsa de valores do Sri Lanka.



Trabalho relacionado	As ações de mais de uma empresa foram avaliadas?	Variáveis Macroeconômicas foram usadas?	Topologia da rede neural usada	Objetivo da pesquisa
Nelson, Pereira e Oliveira (2017)	Sim, ações BOVA11, BBDC4, CIEL3, ITUB4 e PETR4	Não	LSTM (Long Short Term Memory)	Aplicabilidade de redes neurais recorrentes, sobre o problema da previsão de movimentos de preços de mercado de ações.
Xiong e Lu (2017)	Sim, quatro ações do mercado principal quanto no mercado corporativo de crescimento no setor de software e serviços de informação	Não	Híbrido, neste estudo foi utilizado o ARIMA e o Multilayer Perceptron	Prever a tendência do índice de preços das ações
Lam (2004)	Sim, informações de 364 empresas da S&P no período de 1985–1995	Sim	Multilayer Perceptron	Investigar a capacidade de redes neurais, com retropropagação para integrar análise fundamental e técnica para previsão de ações.
Mingyue, Cheng e Yu (2016)	Não, neste estudo foi avaliado as ações do Nikkei 225	Não	Multilayer Perceptron	Prever a direção do índice NIKKEI 225.
Oriani e Coelho (2016)	Sim, neste estudo foram utilizados dados das ações ABEV3, BBAS, CCRO3, PETR4 e VALE5	Não	Multilayer Perceptron	Avaliar o impacto de indicadores técnicos na qualidade das previsões de preço de fechamento de ações baseadas em redes neurais.
Este Trabalho	Sim, as ações com maior representatividade e no mercado de ações da B3	Sim	LSTM	Avaliar o uso de variáveis macroeconômicas na previsão do índice Bovespa com RNAs do tipo LSTM.

**Fonte:** Elaborado pelo autor

Sobre a influência de fatores macroeconômicos, a pesquisa de Sousa (2016) provê evidência de que variáveis macroeconômicas ou macro-financeiras domésticas e internacionais podem prever o retorno (lucro) de ações em mercados emergentes, neste caso o objeto de estudo foram os BRICS fundamentado pelas seguintes razões: Primeiro este bloco econômico representa os mercados com maior taxa de crescimento nas últimas décadas, o segundo motivo é a representatividade econômica deste mercado que corresponde a cerca de 40% da população mundial e aproximadamente 20% do PIB mundial. O terceiro motivo é o quadro institucional econômico para a política monetária cuja meta é estabelecida com base na inflação, deixando as decisões econômicas mais previsíveis. Por fim, políticas macroeconômicas mais sólidas e uma inflação baixa e estável impulsionaram a eficácia da política monetária explícita em um fortalecimento dos mercados internos

através de mais ativos em moeda nacional e condições de empréstimo mais acessíveis longe de ativos estrangeiros.

Sousa (2016) provê evidência sobre a previsibilidade do retorno de ações no uso de variáveis domésticas macroeconômicas, em especial o PIB para o Brasil e China. No entanto outras variáveis como a inflação e taxas de juros mostraram pouco ou nenhuma influência no retorno de ações. A pesquisa do autor é relevante e explora a correlação entre o mercado de ações e variáveis macroeconômicas, porém neste contexto as redes neurais não foram exploradas.

No tocante a aplicação de redes neurais para a previsão da direção do preço de uma ação Oliveira, Nobre e Zárate (2013) desenvolveram uma rede neural capaz de prever a direção da ação PETR4 que faz uso de variáveis macroeconômicas e outras variáveis técnicas. A metodologia adotada de entendimento do domínio do problema, seleção e análise de amostras, pré-processamento das entradas e modelagem e predição se mostrou adequada para o objetivo da pesquisa.

Os autores apresentaram uma metodologia para o experimento e também explicaram como a quantidade de neurônios foi calculada, o que está de acordo com as preocupações de Tkáč e Verner (2016) sobre a descrição da topologia da rede neural projetada. As fontes das variáveis também podem ser verificadas por futuros pesquisadores o que valida a possibilidade de replicação do experimento.

O modelo de rede neural adotado por Oliveira, Nobre e Zárate (2013) foi do tipo feed-forward multilayer perceptron com a seguinte estrutura: 3 camadas, incluindo uma camada de entrada, uma camada oculta e uma camada de saída. O número de neurônios na camada de entrada é determinado pelo tamanho da janela (período em meses) multiplicado pelo número de séries no conjunto de amostras. O número de neurônios na camada oculta foi determinado pela equação:  $2\sqrt{(m+2)N}$  onde  $m$  e  $n$  representam o número de neurônios da saída e da camada de entrada, respectivamente. Para treinamento e validação do algoritmo foi utilizada retropropagação resiliente, uma adaptação do algoritmo padrão de retropropagação.

No entanto, o experimento de Oliveira, Nobre e Zárate (2013) foi realizado somente com uma ação, a PETR4 que corresponde a empresa Petrobrás. Poderiam ter sido usadas mais fontes de dados, outras ações ou até mesmo o Ibovespa para uma avaliação mais completa do método e das variáveis escolhidas.

Outros estudos que merecem destaque e que são relacionados ao mercado de ações brasileiro são os de Silva, Castilho, Pereira e Brandão (2014), Oriani e Coelho (2016) e o de Nelson, Pereira e Oliveira (2017).

A pesquisa de Oriani e Coelho (2016) avaliou o impacto de diferentes indicadores técnicos na previsão de preços de fechamento de ações. A primeira etapa do trabalho foi a escolha de cinco ativos de empresas de diferentes segmentos de mercado, com boa representatividade no índice Ibovespa e volume negociado elevado e sobre estes ativos foi os seguintes indicadores técnicos foram calculados: Índice direcional de média (ADX) utilizado para quantificar a força ou fraqueza de uma tendência observada no mercado, média móvel simples (SMA), média móvel exponencial simples (EMA), média móvel exponencial dupla (DEMA), média exponencial tripla (TEMA) e média móvel ponderada (WMA), são adaptações do indicador SMA utilizados para considerar o ruído e a volatilidade do mercado, outro indicador baseado em médias exponenciais é o KAMA ou adaptação de médias exponenciais de Kaufmann que mostra a oscilação detalhada perto dos valores reais da série temporal, o indicador Momentum demonstra a diferença entre o preço de fechamento atual e o passado, a taxa de variação (ROC) da a diferença percentual no preço de fechamento, o índice de força relativa (RSI) mostra a fraqueza de uma tendência ou ativo considerando os últimos preços de fechamento. Por fim o indicador TRIX mostra a taxa de variação percentual de uma média tripla exponencialmente móvel de um determinado preço de fechamento de ações enquanto o indicador Williams %R mostra o preço de fechamento de uma ação associada aos altos e baixos preços observados aos últimos  $n$  instantes da série temporal.

Para avaliar o impacto de cada indicador uma rede neural do tipo multilayer perceptron foi desenvolvida com o objetivo de prever o preço de fechamento de D+1. Os autores justificam a escolha deste algoritmo por causa de sua capacidade de aproximar funções e identificar padrões. A conclusão do experimento provê evidência empírica do sobre a melhoria da acurácia da previsão dos preços de fechamento das ações. No entanto não era objetivo deste estudo avaliar o impacto de variáveis macroeconômicas.

No experimento de Silva, Castilho, Pereira e Brandão (2014) a RNA do tipo multilayer perceptron foi modelada com o objetivo prever a oscilação positiva dos preços das ações em curtos períodos de tempo (5, 10 ou 15 minutos), para isso foram

utilizados variáveis de análise técnica calculadas com base nos gráficos de velas correspondentes a algumas ações comercializadas no mercado brasileiro.

O trabalho de Nelson, Pereira e Oliveira (2017) também teve por objetivo prever a oscilação positiva em um curto espaço de tempo (15 minutos), para isso foi utilizado um outro tipo de topologia de rede neural a LSTM. As variáveis utilizadas também são específicas da análise técnica.

Sobre o contexto internacional, outros trabalhos que também estão relacionados a esta pesquisa e utilizaram dados de outros mercados de ações de outros países são os de Lam (2004), Altinbas e Biskin (2015), Mingyue, Cheng e Yu (2016), Gunduz, Yaslan e Cataltepe (2017), Samarawickrama e Fernando (2017), Xiong e Lu (2017).

O trabalho de Lam (2004), teve por objetivo investigar a capacidade de redes neurais, especificamente, o algoritmo de retropropagação para integrar análise fundamental e técnica para previsão de desempenho financeiro. A taxa de retorno do patrimônio líquido é usada como a variável a ser prevista. Para o desenvolvimento desta pesquisa foram utilizados dados macroeconômicos e também de análise técnica, a topologia utilizada foi a multilayer perceptron. A quantidade de dados corresponde a uma parcela significativa das empresas que compõem o índice S&P, neste estudo foram utilizados os dados de 364 empresas. Apesar dos resultados positivos e do experimento o estudo utilizou variáveis macroeconômicos relevantes ao mercado de ações estadounidense que é diferente do brasileiro.

Mingyue, Cheng e Yu (2016) utilizaram uma rede neural do tipo multilayer perceptron e a retropropagação com o objetivo de prever a direção do índice diário do mercado de ações do Japão, o Nikkei 225. Variáveis da análise técnica foram usadas neste experimento. O algoritmo genético, ou GA, também foi utilizado para otimizar os pesos iniciais e o viés do modelo da RNA.

Samarawickrama e Fernando (2017) e Xiong e Lu (2017) utilizaram de uma abordagem híbrida em suas pesquisas. A pesquisa de Samarawickrama e Fernando (2017) utilizou dados da bolsa de valores do Sri Lanka e fez uma comparação entre diversas topologias de redes neurais. Xiong e Lu (2017) desenvolveu sua pesquisa com dados da bolsa de valores chinesa e o algoritmo desenvolvido era híbrido com elementos de uma rede neural do tipo multilayer perceptron e o ARIMA.

O estudo de Altinbas e Biskin (2015) não informa como a topologia da rede neural foi desenhada. Apenas como ela foi aplicada. Para isso, foi usado as variáveis mais relevantes ao contexto do mercado de ações turco. O método utilizado foi uma rede neural do tipo Multilayer Perceptron.

Gunduz, Yaslan e Cataltepe (2017) aplicam redes neurais para a previsão da direção de 100 ações da bolsa de Istambul em uma janela de tempo de 1 hora.

O tipo de rede neural usado foi a convolucional. A principal diferença entre uma rede neural artificial e uma rede neural convolucional é a conexão entre as camadas. Na rede neural convolucional, cada parte local (campo receptivo) das entradas é conectada a apenas um neurônio, enquanto as entradas na rede neural artificial são totalmente conectadas aos neurônios na próxima camada. Em cada camada da rede convolucional, a operação de convolução é feita através da aplicação de filtros de tamanhos diferentes nas entradas.

O algoritmo desenvolvido tem um total de 7 camadas com 1 camada de entrada, 2 camadas de convolução paralelas, 1 camada de mesclagem, 1 camada de convolução, 1 camada totalmente conectada (densa) e 1 camada de saída.

O experimento de Gunduz, Yaslan e Cataltepe (2017) usa uma janela de tempo diferente do experimento Oliveira, Nobre e Zárate (2013), a topologia da rede neural utilizada também. No entanto, as variáveis utilizadas não são relativas a fatores macroeconômicos.

Os artigos avaliados apresentam topologias e algoritmos de aprendizagem diferentes de redes neurais mostrando a sua eficácia no mercado financeiro. No entanto, os estudos apresentados utilizam variáveis correspondentes as características dos mercados locais e um foco muito grande em variáveis de análise técnica, o que evidencia a necessidade de um estudo mais detalhado sobre o uso de redes neurais e variáveis macroeconômicas. Para o objetivo deste estudo será utilizado o contexto do mercado de ações brasileiro.

Nesta seção foram abordados os estudos sobre aplicação de séries temporais que fundamentam este trabalho, especialmente a pesquisa de Oriani e Coelho (2016), que serviu de “benchmark” para nossos estudos, por tratar séries que apresentam características semelhantes e métodos específicos para avaliar o impacto das variáveis técnicas em redes neurais.

## 4 MÉTODO DE PESQUISA

### 4.1 Introdução

O objetivo deste capítulo é mostrar o tipo de pesquisa realizada neste trabalho, a amostra utilizada e o período selecionado para o experimento. Além disso são mostrados os dados coletados, o algoritmo desenvolvido e o método de avaliação a ser usado para posterior análise e interpretação.

### 4.2 Método de Pesquisa

Para Lacerda et al. (2013) o enquadramento metodológico de uma pesquisa não deve ser compreendido como um ato burocrático. Ele consiste em escolher e justificar um método de pesquisa que possibilite: Responder ao problema de pesquisa formulado, ser avaliado pela comunidade científica e evidenciar procedimentos que robusteçam os resultados da pesquisa.

Marconi e Lakatos (2008) definem como método um conjunto de atividades sistemáticas e racionais que com maior segurança e economia, possibilitam alcançar o objetivo - conhecimentos válidos e verdadeiros - traçando o caminho a ser seguido, detectando erros e auxiliando as decisões do cientista.

Com relação aos métodos de pesquisa existentes, o presente estudo pode seguir a abordagem de Design Science Research conforme proposta de Lacerda et al. (2013). O método em si diferencia-se de um estudo de pesquisa de ação tradicional cujas principais características são as elencadas por Coughlan e Coughlan (2002): Primeiro, o pesquisador toma ação, depois ele deve envolver dois objetivos: Resolver um problema e contribuir para a ciência. Por fim requer a cooperação entre pesquisadores e os participantes da pesquisa. Lacerda et al. (2013) mostra a diferença entre os métodos ao mostrar o papel do pesquisador em avaliar seu artefato, possibilidade que não é considerada na pesquisa-ação tradicional. Não há, também, necessidade explícita de contribuir para uma teoria de cunho prescritivo, pois o foco principal é a utilidade no caso singular em observação.

Segundo Lacerda et al. (2013) não há consenso, em particular, sobre as fronteiras entre a Pesquisa-Ação e a Design Science Research na literatura. As diferenças entre os métodos de pesquisa podem ser observadas no quadro abaixo.

**Quadro 2** - Comparativo entre a Design Science Research, o Estudo de Caso e a Pesquisa-Ação

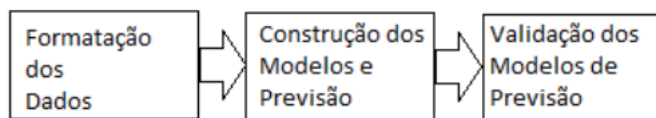
Características	<i>Design Science Research</i>	Estudo de Caso tradicional	Pesquisa-Ação tradicional
Objetivos	Desenvolver artefatos que permitam soluções satisfatórias aos problemas práticos. Prescrever e Projetar	Auxiliar na compreensão de fenômenos sociais complexos. Explorar, Descrever e Explicar	Resolver ou explicar problemas de um determinado sistema gerando conhecimento para a prática e para a teoria. Explorar, Descrever e Explicar
Principais Atividades	<ul style="list-style-type: none"> <li>• Conscientizar</li> <li>• Sugerir</li> <li>• Desenvolver</li> <li>• Avaliar</li> <li>• Concluir</li> </ul>	<ul style="list-style-type: none"> <li>• Definir Estrutura Conceitual</li> <li>• Planejar o(s) caso(s)</li> <li>• Conduzir Piloto</li> <li>• Coletar Dados</li> <li>• Analisar Dados</li> <li>• Gerar Relatório</li> </ul> Miguel (2007, p. 221)	<ul style="list-style-type: none"> <li>• Planejar a Ação</li> <li>• Coletar Dados</li> <li>• Analisar dados e Planejar ações</li> <li>• Implementar Ações</li> <li>• Avaliar Resultados</li> <li>• Monitorar (Contínuo)</li> </ul> Turrión e Mello (2010)
Resultados	Artefatos (Constructos, Modelos, Métodos, Instanciações)	Constructos Hipóteses Descrições Explicações	Constructos Hipóteses Descrições Explicações Ações
Tipo de Conhecimento	Como as coisas deveriam ser	Como as coisas são ou como se comportam.	Como as coisas são ou como se comportam.
Papel do Pesquisador	Construtor e Avaliador do Artefato	Observador	Múltiplo, em função do Tipo de Pesquisa-Ação
Base Empírica	Não obrigatória	Obrigatória	Obrigatória
Colaboração Pesquisador-Pesquisado	Não obrigatória	Não obrigatória	Obrigatória
Implementação	Não obrigatória	Não se Aplica	Obrigatória
Avaliação dos Resultados	Aplicações Simulações Experimentos	Confronto com a Teoria	Confronto com a Teoria
Abordagem	Qualitativa e/ou Quantitativa	Qualitativa	Qualitativa

**Fonte:** Lacerda et al. (2013)

A natureza do problema de pesquisa deste trabalho e a comparação apresentada no Quadro 2 acima indicam o método de Design Science Research como a melhor alternativa.

#### 4.3 Configurações experimentais

Esta seção descreve as etapas do trabalho proposto e a implementação realizada detalhando os processos. Na primeira etapa os dados foram coletados e formatados para utilização. Depois os modelos de previsão são aplicados e os critérios de medição e validação do experimento são aferidos. O diagrama apresentado na figura 12 ilustra o processo para avaliação do experimento:

**Figura 12** - Diagrama do processo experimental a ser aplicado

**Fonte:** Elaborado pelo autor

#### 4.4 Conjunto de dados

A primeira etapa na seleção de variáveis de uma rede neural consiste na decisão do horizonte temporal a ser analisado. Algumas possibilidades são previsões anuais, mensais, semanais e diários. Para este experimento os dados coletados correspondem a periodicidade diária.

O conjunto de dados utilizado neste trabalho consiste de dados reais de negociações na B3. Os ativos VALE3, PETR4, ITUB4, BBDC4 e BBSA3 foram selecionados porque tinham a maior representatividade na composição do Ibovespa em Dezembro de 2018. Também foi selecionado o ativo BOVA11, que pode ser definido como um *exchange traded fund* ou ETF. Um ETF é uma unidade negociável na bolsa de valores da mesma forma que uma ação, mas neste caso representando o comportamento de um índice, neste caso o Ibovespa.

Os dados de entrada utilizados neste trabalho correspondem ao intervalo de 03/08/2008 e 31/12/2018 totalizando 2490 registros. Estes dados foram extraídos do arquivo de movimentações disponibilizado pela B3. A figura 13 mostra a evolução do preço de fechamento durante o período de estudo.

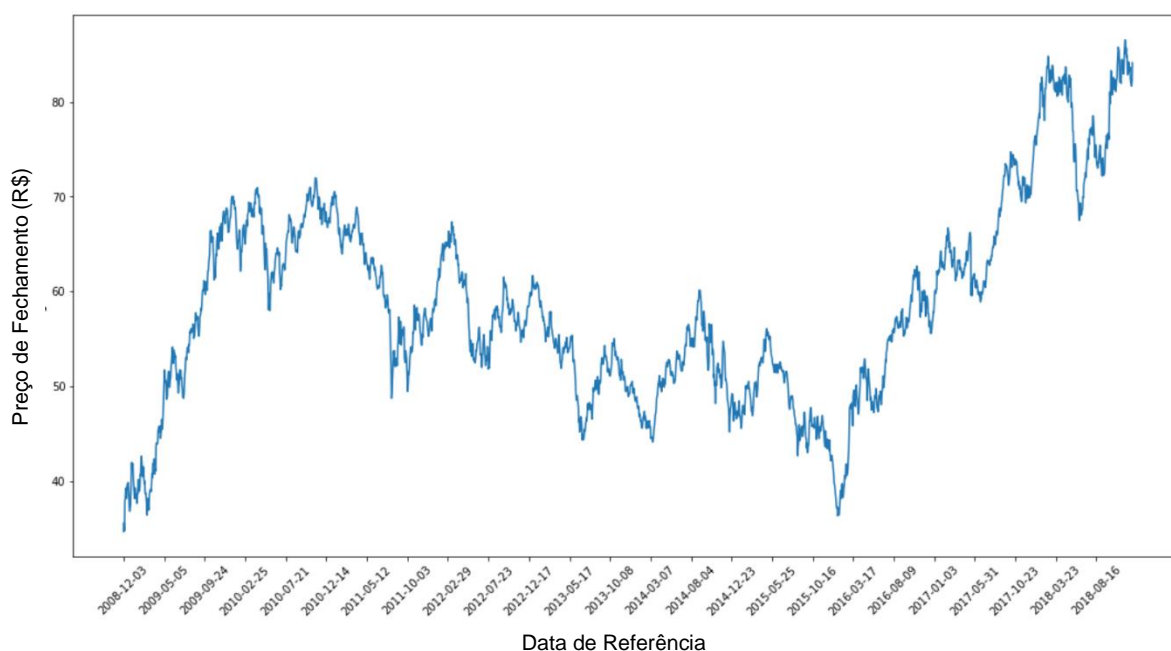
**Tabela 01** - Subconjunto da amostra com os 5 primeiros registros do arquivo – BOVA11

DATA	ABERTURA	MÍNIMA	MÁXIMA	FECHAMENTO	VOLUME
03 DEZ 2008	34,73	33,90	35,40	35,31	330.600
04 DEZ 2008	35,80	35,15	36,00	35,70	128.400
05 DEZ 2008	34,51	34,11	35,35	35,35	193.700
08 DEZ 2008	37,35	36,95	38,46	38,46	103.100
09 DEZ 2008	38,58	37,54	38,82	37,54	274.500

**Fonte:** Elaborado pelo autor



**Figura 13** - Representação da evolução do preço de fechamento do ETF BOVA11 no período de 03/12/2008 a 31/12/2018



**Fonte:** Elaborado pelo autor

Keiel e Bender (2017) afirmam que a utilização de variáveis exógenas na predição apresenta maior eficácia do que modelos sem este tipo de variável. Kaastra e Boyd (1996) citam em seu estudo que os atributos utilizados na previsão de preços através de uma rede neural podem ser classificados de 3 maneiras: Primeiro há os indicadores técnicos tais como: Preço de abertura, preço de fechamento, preço máximo, preço mínimo, volume de negociações, etc. Os dados contábeis (balanços patrimoniais e projeções, por exemplo) e os dados de natureza macroeconômica (por exemplo, taxa de juros e índice IGPM).

Os dados contábeis e os indicadores técnicos são disponibilizados em diversos canais e podem facilmente ser acessados, no entanto, a respeito dos dados macroeconômicos Kaastra e Boyd (1996) comentam que deve ser verificado se a série temporal correspondente é consistente ao longo do tempo, se são sujeitas a revisões periódicas e se são divulgadas com atraso.

Kaastra e Boyd (1996) também sugerem alguns procedimentos para minimizar ruídos que prejudicam a identificação de tendências e adequar a distribuição de ocorrências das variáveis às assimetrias. Os autores indicam: O uso do logaritmo da variável para reduzir distribuições de dados onde o volume é classificado de forma

decrecente, redução das variáveis redundantes e no retreinamento periódico ou recalibragem da rede de forma que ela esteja mais adequada a diferentes condições de mercado ou ciclos econômicos.

O estudo de Kaastra e Boyd (1996) também recomenda o teste de múltiplas combinações de indicadores. Esta recomendação foi adotada por muitos pesquisadores, e ao longo do tempo um outro comportamento que é a utilização de variáveis que poderiam descrever a anomalia do preço momentâneo dos ativos financeiros pôde ser observada, evento identificado como *anomalia de efeito momento* que caracteriza um grau de previsibilidade ao mercado acionário.

O efeito momento foi identificado primeiramente em outro experimento pelos pesquisadores Jegadeesh e Titman (1993) através de um experimento que verificou o desempenho de 2 carteiras durante seis meses para a formação das carteiras vencedoras e perdedoras, seguidos de trinta e seis meses nos quais os retornos destas carteiras são comparados com a média do mercado. De acordo com a hipótese do mercado eficiente de Fama (1964) as carteiras vencedoras deveriam ter seu retorno sempre superior ao das carteiras perdedoras, porém o resultado observado foi o oposto. Os excessos de retorno das carteiras perdedoras foram maiores que os excessos das vencedoras. Fama (1997) reconheceu o efeito momento como um parâmetro de eficiência de mercado contrário a sua hipótese de mercado eficiente.

Murphy (1999) descreve que uma vez que uma tendência é identificada nos preços de um ativo é provável que ela se mantenha no curto prazo ao invés de ser revertida. Conclui-se então que os atributos do vetor de entrada das redes neurais devem ser escolhidos de forma a capturar a anomalia do efeito momento dos preços dos ativos.

### **Quadro 03 – Descrição das variáveis técnicas**

<b>Variável</b>	<b>Descrição</b>
Abertura	Preço de abertura do ativo no pregão.
Fechamento	Preço de fechamento do ativo no pregão.
Máximo	Preço máximo do ativo no pregão.
Mínimo	Preço mínimo do ativo no pregão.
Volume	Volume de transações do ativo.

**Fonte:** Elaborado pelo autor

**Quadro 04** – Descrição das variáveis contábeis e macroeconômicas

<b>Variável</b>	<b>Descrição</b>
IGPM	Índice Geral de Preços do Mercado, indicador que registra a inflação de preços de matérias-primas agrícolas e industriais, bens e serviços finais. É medido pela Fundação Getúlio Vargas (FGV).
IPCA	Índice Nacional de Preços ao Consumidor Amplo é o índice oficial do governo federal brasileiro para medição de metas inflacionárias. É medido pelo Instituto Brasileiro de Geografia e Estatística (IBGE).
Taxa SELIC	Taxa básica de juros da economia brasileira, refere-se à taxa de juros apurada nas operações de empréstimos de um dia entre as instituições financeiras que utilizam títulos públicos federais como garantia. É o principal instrumento de política monetária utilizado pelo Banco Central para controle da inflação.
Taxa CDI	Taxa que representa o lastro de operações do mercado interbancário, é mantida pela CETIP que foi incorporada a B3.
Índice de Confiança do Consumidor	Índice Geral de Expectativa do Consumidor, mede a confiança dos consumidores.
Índice de Expectativas Futuras	Indicador de acompanhamento das expectativas atuais e futuras de consumidores e de empresários em relação à situação econômica, é mantido pelo Banco Central.
Índice de Commodities de Energia Brasileiras	Indicador construído com base no preço das commodities agrícolas, metálicas e energéticas, é mantido pelo Banco Central.
EMBI – Risco Brasil	O EMBI é um índice baseado nos bônus (títulos de dívida) emitidos pelos países emergentes. Mostra os retornos financeiros obtidos a cada dia por uma carteira selecionada de títulos desses países.
Índice de Atividade Econômica do Banco Central IBC-BR	O Índice de Atividade Econômica do Banco Central do Brasil (IBC-Br) é um indicador criado para tentar antecipar o resultado do Produto Interno Bruto (PIB) do país, constituindo-se em um parâmetro preliminar da evolução da atividade econômica brasileira.
Fechamento IBOVESPA	Índice da Bolsa de Valores de São Paulo (Ibovespa).
PIB Real	Produto Interno Bruto (PIB) real a preços de mercado. Variação percentual em relação ao mesmo período do ano anterior.
Cotação do Dólar	Taxa de câmbio R\$/US\$ comercial (valor de compra).

**Fonte:** Elaborado pelo autor

Com o objetivo de validar o estudo proposto por Oliveira, Nobre e Zárate (2013) e a influência do contexto macroeconômico brasileiro proposto por Garcia e Vasconcellos (2002) as seguintes variáveis foram selecionadas: IGPM, IPCA Taxa Selic, Taxa CDI, o Índice de Confiança do Consumidor, Índice de Expectativas Futuras

e o Índice de Commodities de Energia Brasileira. Sousa (2016) demonstrou em seu estudo que o PIB também pode influenciar diretamente os resultados do mercado de ações, por este motivo o índice de atividade econômica do Banco Central IBC-BR e o PIB Real também foram selecionados. Por fim também foi acrescido a este estudo o índice EMBI, o indicador IBOVESPA e a cotação de fechamento do dólar norte-americano.

As tabelas 02, 03, 04, 05 e 06 mostram a correlação linear de Pearson que mede o grau da correlação linear entre duas variáveis quantitativas. com valores situados ente 1 (correlação forte direta) e -1 (correlação forte inversa) que reflete a intensidade de uma relação linear entre dois conjuntos de dados.

**Tabela 02** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo BOVA11

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.997399
Variação	0.045480
Variação %	0.047528
Preço Mínimo	0.998958
Preço Máximo	0.998915
IGPM	0.035362
IPCA	-0.306947
SELIC	-0.595223
Índice Nacional de Expectativas do Consumidor	0.155241
Índice de Expectativas Futuras	0.302823
Índice de Commodities Energéticas Brasileira	0.223938
EMBI – Risco Brasil	-0.468195
Índice de Atividade Econômica do Banco Central IBC-BR	-0.035351
Fechamento IBOVESPA	0.997251
PIB Real	0.477463
Cotação do Dólar	0.060460

**Fonte:** Elaborado pelo autor

**Tabela 03** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo VALE3

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.998409
Variação	0.034352
Variação %	0.021888
Preço Mínimo	0.999356
Preço Máximo	0.999321
IGPM	0.084208
IPCA	-0.187847
SELIC	-0.617610
Índice Nacional de Expectativas do Consumidor	0.657721
Índice de Expectativas Futuras	0.674615
Índice de Commodities Energéticas Brasileira	-0.348304
EMBI – Risco Brasil	-0.620029
Índice de Atividade Econômica do Banco Central IBC-BR	-0.011042
Fechamento IBOVESPA	0.663470
PIB Real	0.783476
Cotação do Dólar	-0.489359

**Fonte:** Elaborado pelo autor

**Tabela 04** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo PETR4

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.998461
Variação	0.033101
Variação %	0.023251
Preço Mínimo	0.999378
Preço Máximo	0.999274
IGPM	-0.153665

<b>Variável</b>	<b>Pearson</b>
IPCA	-0.250930
SELIC	-0.437645
Índice Nacional de Expectativas do Consumidor	0.764783
Índice de Expectativas Futuras	0.700688
Índice de Commodities Energéticas Brasileira	-0.655543
EMBI – Risco Brasil	-0.468729
Índice de Atividade Econômica do Banco Central IBC-BR	-0.263489
Fechamento IBOVESPA	0.352682
PIB Real	0.643913
Cotação do Dólar	-0.671470

**Fonte:** Elaborado pelo autor

**Tabela 05** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo ITUB4

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.997008
Variação	0.062373
Variação %	0.998780
Preço Mínimo	0.998783
Preço Máximo	0.091396
IGPM	-0.023307
IPCA	-0.241594
SELIC	-0.368637
Índice Nacional de Expectativas do Consumidor	-0.332163
Índice de Expectativas Futuras	-0.273910
Índice de Commodities Energéticas Brasileira	0.656835

<b>Variável</b>	<b>Pearson</b>
EMBI – Risco Brasil	-0.081424
Índice de Atividade Econômica do Banco Central IBC-BR	-0.024981
Fechamento IBOVESPA	0.823612
PIB Real	0.018997
Cotação do Dólar	0.535035

**Fonte:** Elaborado pelo autor

**Tabela 06** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo BBDC4

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.997416
Variação	0.057635
Variação %	0.045373
Preço Mínimo	0.998949
Preço Máximo	0.998928
IGPM	-0.014023
IPCA	-0.193037
SELIC	-0.357730
Índice Nacional de Expectativas do Consumidor	-0.372724
Índice de Expectativas Futuras	-0.355970
Índice de Commodities Energéticas Brasileira	0.702821
EMBI – Risco Brasil	-0.139483
Índice de Atividade Econômica do Banco Central IBC-BR	0.170463
Fechamento IBOVESPA	0.713887
PIB Real	-0.058604
Cotação do Dólar	0.542213

**Fonte:** Elaborado pelo autor

**Tabela 07** – Demonstração da correlação Pearson com a variável alvo, Preço de Fechamento – Ativo BBSA3

<b>Variável Fechamento</b>	
<b>Variável</b>	<b>Pearson</b>
Abertura	0.995872
Variação	0.066322
Variação %	0.050674
Preço Mínimo	0.998353
Preço Máximo	0.998237
IGPM	-0.093718
IPCA	-0.322103
SELIC	-0.536993
Índice Nacional de Expectativas do Consumidor	0.164035
Índice de Expectativas Futuras	0.189515
Índice de Commodities Energéticas Brasileira	0.184781
EMBI – Risco Brasil	-0.514342
Índice de Atividade Econômica do Banco Central IBC-BR	0.097379
Fechamento IBOVESPA	0.908847
PIB Real	0.456151
Cotação do Dólar	0.010375

**Fonte:** Elaborado pelo autor

A análise desta técnica em diferentes ativos possibilitou visualizar as seguintes correlações correspondentes a variáveis macroeconômicas com o preço de fechamento. Sobre o ativo BOVA11 é possível observar a correlação forte entre o fechamento e o Índice IBOVESPA. No ativo VALE3 existe correlação forte entre o preço de fechamento e o PIB Real, o que corrobora o estudo de Sousa (2016) e sua hipótese de influência do PIB Real na previsão do mercado financeiro, outras variáveis macroeconômicas que também podem influenciar o comportamento deste ativo são: O Índice Nacional de Expectativas do Consumidor e o Índice Nacional de Expectativas Futuras. Uma correlação muito parecida pode ser observada no ativo PETR4, no



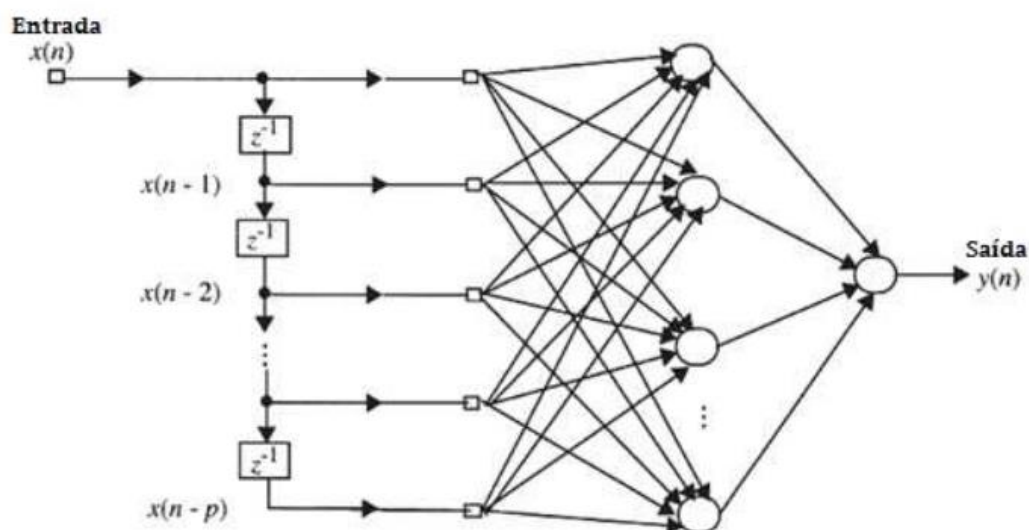
entanto a correlação do preço de fechamento é mais forte com o Índice Nacional de Expectativas do Consumidor e o Índice Nacional de Expectativas Futuras do que o PIB Real. Os demais papeis avaliados ITUB4, BBDC4 e BBSA3 correspondem respectivamente aos bancos Itaú, Bradesco e Banco do Brasil, para estes ativos é possível observar uma correlação forte para a variável de fechamento do IBOVESPA, para o ativo ITUB4 também existe uma correlação média com o Índice de Commodities Energéticas Brasileira.

#### 4.4 Janelas móveis

Haykin (2001) cita que em alguns casos como a previsão de séries temporais o tempo deve ser incorporado como uma dimensão de aprendizado de maneira implícita ou explícita. Segundo Cortez (2007) a previsão de séries temporais pode ser executada através de uma janela deslizante através de uma estrutura de memória de curto prazo formada por uma memória do tipo “linha de atraso com derivações”. Uma linha de atraso  $X$  unidades de atraso com  $(X + 1)$  terminais. Esta técnica consiste no treinamento cíclico da rede neural empregada na previsão correspondente a períodos de verificação (móveis ou deslizantes) ao longo do tempo. Kaastra e Boyd (1996) recomendam a utilização desta técnica de forma a rede neural se adaptar a diferentes condições e fatores de influência mais recentes. A cada amostra lida pela rede neural da série acontece um deslocamento na posição dos valores presentes na entrada da linha de atraso transferindo os valores de uma entrada para sua adjacente menos significativa. Haykin (2001) afirma que esta técnica garante que a ordem de leitura da série preservada, mantendo assim a dimensão de tempo e a dependência entre as amostras.

O termo janela móvel ou janela deslizante é descrito pela forma com que os dados da série são lidos sequencialmente, o último valor é descartado a cada nova amostra que é lida e a manutenção do mesmo número de entradas da rede. A janela pode ser descrita como um vetor com números subsequentes como  $< 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 >$  para um  $Y=10$ . Outra possibilidade é a aplicação em valores sazonais em que o vetor representa o deslocamento na leitura de amostrar como por exemplo no vetor  $< 1, 8, 15, 22, 29 >$  para um  $Y=5$  e um  $A=7$ , onde  $A$  representa a distância de leitura entre as amostras. A figura 14 representa a linha de atraso e a técnica de janelas móveis em combinação com uma rede neural do tipo multilayer perceptron.

**Figura 14** – Representação de uma rede neural artificial com uma linha de atraso



**Fonte:** Adaptado de Haykin (2001).

O emprego desta técnica possibilita utilizar dados macroeconômicos como entradas da rede neural que geralmente são oriundos de pesquisas ou coletas com uma sazonalidade diferente das séries temporais diárias dos ativos de ações.

#### 4.5 Python

O Python foi criado por criada por Guido Van Rossum em 1991, esta linguagem é do tipo interpretada e orientada a objetos. Este ambiente de desenvolvimento escolhido para a implementação proposta por apresentar bibliotecas específicas que podem ser úteis neste projeto, e são utilizadas para o processamento de grandes volumes de dados, otimização, álgebra linear e integração com frameworks de *deep learning*.

#### 4.6 Arquitetura das Redes Neurais e configurações experimentais

Zhang (2001) mostrou em sua pesquisa que os principais parâmetros na utilização de redes neurais para a previsão de séries temporais são: O número de nós de entra e das camadas ocultas da rede neural e o tamanho da amostra a ser avaliada. As redes neurais não necessariamente necessitam de grandes amostras de dados para obter um desempenho satisfatório.

Para este estudo a quantidade de nós nas camadas de entrada e oculta demonstrou maior relevância quanto comparado com o tamanho da amostra uma vez

que os dados coletados cobrem todo o período desde o início da comercialização do ETF BOVA11 na B3. Zhang (2001) salienta a importância para a escolha do número de nós da rede nas camadas ocultas pelo seguinte motivo, uma rede neural com poucos nós na camada oculta pode ter sua capacidade de aprendizado reduzida enquanto o excesso de nós pode levar a uma situação de sobre-treinamento ou treinamento em excesso (*overfitting*) em que o modelo perde consideravelmente sua capacidade de generalização embora pareça acurado em uma situação de treinamento. Cortez (2007) também identificou em seu estudo que a quantidade de nós de uma rede neural é a característica que mais afeta a eficácia da previsão.

Weigend e Gershefeld (1994) afirmam que não é possível construir um algoritmo de previsão que seja universal e que seja adequado para qualquer tipo de série temporal. Baseado nesta afirmação o programa desenvolvido foi testado com diversas configurações e combinações de camadas e variáveis.

A rede neural do tipo LSTM foi selecionada para este trabalho por ser frequentemente utilizada em pesquisas similares recentes (Nelson, Pereira e Oliveira 2017 e Samarawickrama e Fernando 2017) e possibilitar a implementação da técnica de janelamento. A rede foi modelada em 3 camadas, uma de entrada, uma camada oculta e outra de saída. Esta topologia segue a recomendação de Fausett (1994) que afirma que o número de camadas ocultas de uma rede neural pode variar e que em alguns casos uma única camada é suficiente.

Os dados que correspondem ao período entre 03 de dezembro de 2008 e 28 de dezembro de 2018 para cada ativo foram divididos em treino e teste. Para o treinamento da rede foram selecionados 70% e o restante (30%) para validação. Como entrada da rede neural as seguintes situações foram testadas:

- Rede Neural LSTM considerando somente variáveis técnicas.
- Rede Neural LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.
- Rede Neural LSTM considerando todas as variáveis.

O algoritmo implementado foi desenvolvido com a capacidade de variar a janela de entrada dos dados entre os dias D-1, D-15 e D-30. Por exemplo, a função de janelamento D-30 utiliza como entrada os 30 (trinta) dias anteriores como input para

a previsão. Desta maneira foi possível testar a entrada da rede considerando diferentes horizontes temporais.

A rede foi treinada utilizando o método de otimização Adam seguindo a recomendação de Kingma e Ba (2015) que demonstraram em seu estudo a sua eficiência quando comparado a outros métodos de otimização estocástica. Este algoritmo tem o gradiente com uma taxa de aprendizado adaptativa que tem por objetivo diminuir a entropia cruzada da rede, o que o torna apropriado para objetivos não estacionários e problemas com muito ruído. O modelo também utilizou a biblioteca Tensorflow que funciona baseada em outra interface da biblioteca Keras.

Por fim outro parâmetro que merece destaque foi o número de épocas, termo que significa o número máximo de iterações que cada rede deveria fazer, não foi encontrado na pesquisa bibliográfica uma recomendação ou modelo para estabelecer este parâmetro optou-se por avaliar o número de 10000 épocas por rede neural testada.

#### 4.7 RMSE

Os modelos foram avaliados através da fórmula RMSE ou raiz do erro médio quadrático, comumente utilizado para expressar a acurácia de um modelo. O RMSE é calculado pela seguinte fórmula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (11)$$

Onde  $P_i$  é o valor previsto,  $O_i$  o valor observado e  $n$  é o número total de observações. Segundo Oliveira, Nobre e Zárate (2013) o RMSE pode ser interpretado como indicador para o poder preditivo ou acordo de previsão. Quanto mais baixo o valor RMSE mais acurada a previsão de acordo com os dados reais.

## 5 ANÁLISE DOS RESULTADOS

Este capítulo tem por objetivo apresentar os resultados aferidos nesta pesquisa. Primeiramente é apresentado os resultados da rede neural do tipo LSTM com as configurações testadas, o gráfico apresentado corresponde ao resultado mais acurado, ou seja, com o menor RMSE. Por fim estes são analisados e comparados com o intuito de validar se os objetivos iniciais da pesquisa foram atendidos.

A rede LSTM desenvolvida correspondente ao ativo BOVA 11 apresentou os resultados apresentados na tabela 8. Cabe a ressalva de que para este caso foi considerado o PIB Real como variável macroeconômica com maior correlação. A variável Fechamento Ibovespa não foi analisada uma vez que este ativo reflete o comportamento do índice.

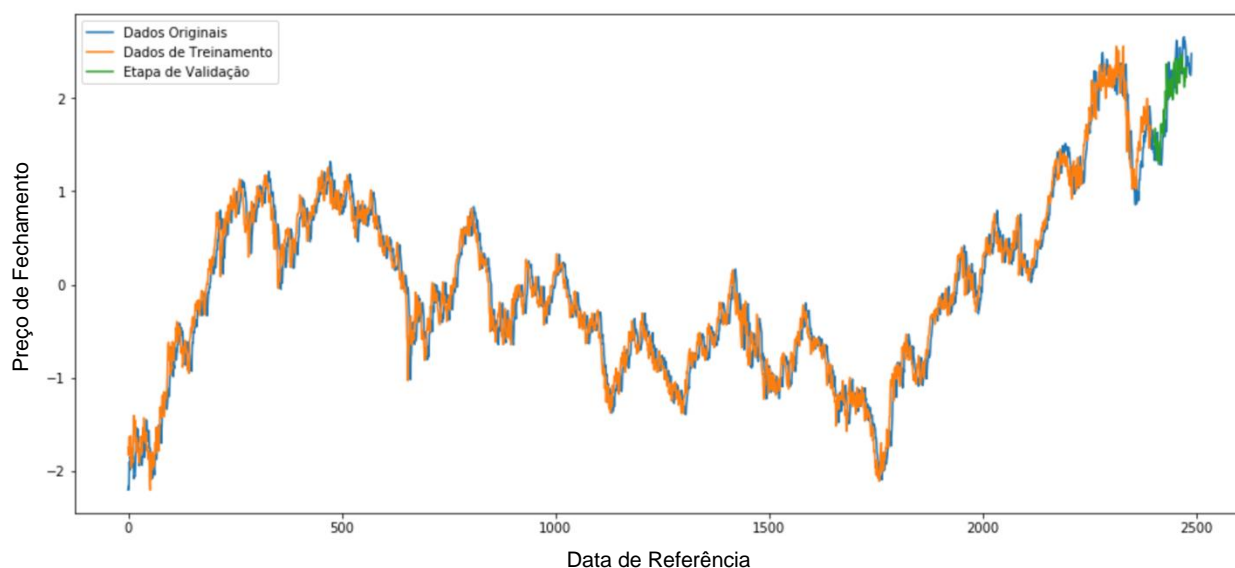
**Tabela 08** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BOVA11

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
PIB Real	LSTM considerando somente variáveis técnicas.	D-1	0,24523
		D-15	0,25782
		D-30	0,25242
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0,97643
		D-15	0,97244
		D-30	0,98121
	LSTM considerando todas as variáveis.	D-1	0,99233
		D-15	0,99245
		D-30	0,99937

**Fonte:** Elaborado pelo autor

O gráfico de saída para o ativo BOVA11 pode ser visto na figura 15.

**Figura 15** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BOVA11



**Fonte:** Elaborado pelo autor

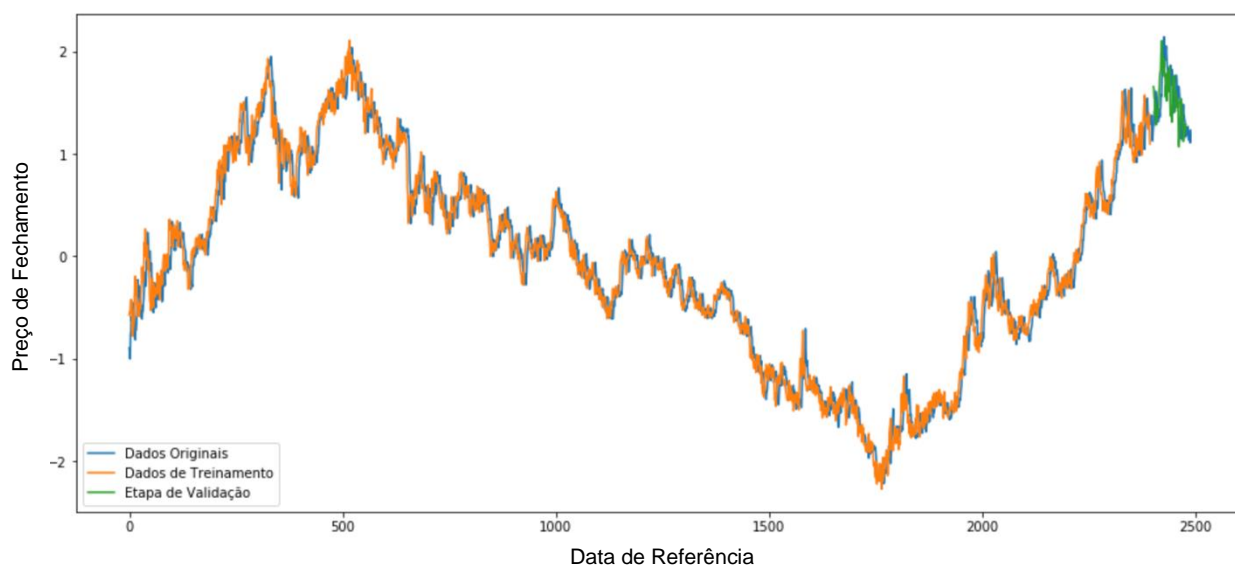
A tabela 09 a seguir mostra os resultados para o ativo VALE3, o gráfico correspondente pode ser visualizado na figura 16.

**Tabela 09** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo VALE3

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
PIB Real	LSTM considerando somente variáveis técnicas.	D-1	0,15442
		D-15	0,14566
		D-30	0,14224
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0,96445
		D-15	0,96882
		D-30	0,96824
	LSTM considerando todas as variáveis.	D-1	0,97314
		D-15	0,97364
		D-30	0,97365

**Fonte:** Elaborado pelo autor

**Figura 16** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo VALE3



**Fonte:** Elaborado pelo autor

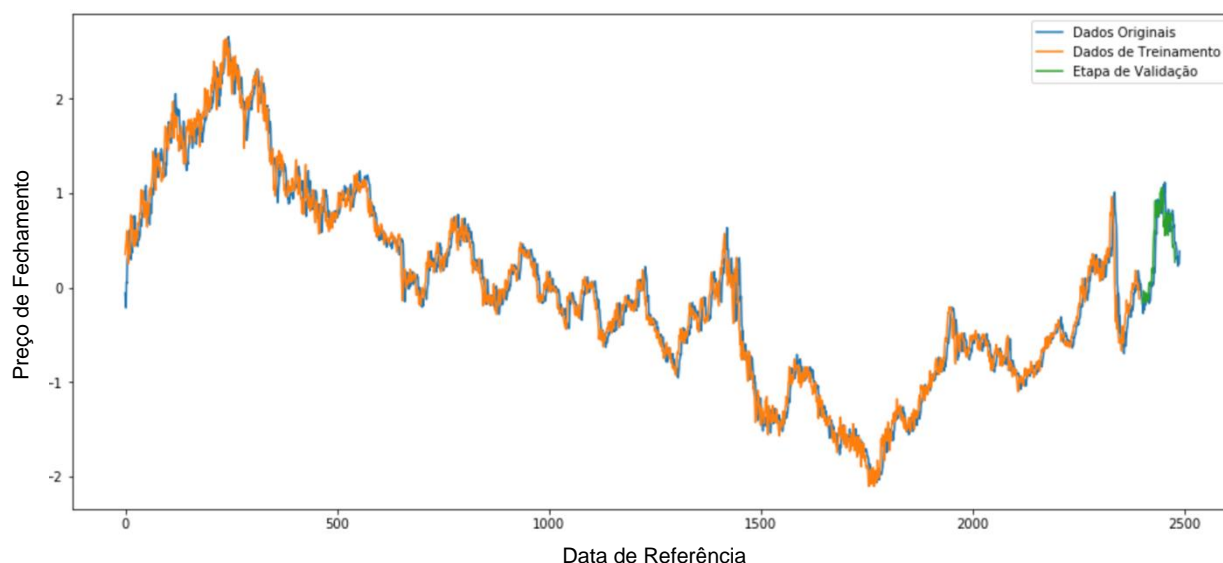
A tabela 10 a seguir mostra os resultados para o ativo PETR4, o gráfico correspondente pode ser visualizado na figura 17. Neste caso, 197 observações foram excluídas do experimento por apresentar valores nulos (NULL) na coluna “Índice Nacional de Expectativas do Consumidor”.

**Tabela 10** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo PETR4.

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
Índice Nacional de Expectativas do Consumidor	LSTM considerando somente variáveis técnicas.	D-1	0,12245
		D-15	0,13422
		D-30	0,13667
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0.69256
		D-15	0.69455
		D-30	0.69779
	LSTM considerando todas as variáveis.	D-1	0,71424
		D-15	0,71345
		D-30	0,71224

**Fonte:** Elaborado pelo autor

**Figura 17** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo PETR4



**Fonte:** Elaborado pelo autor

A tabela 11 mostra os resultados para o ativo ITUB4, o gráfico correspondente pode ser visualizado na figura 18.

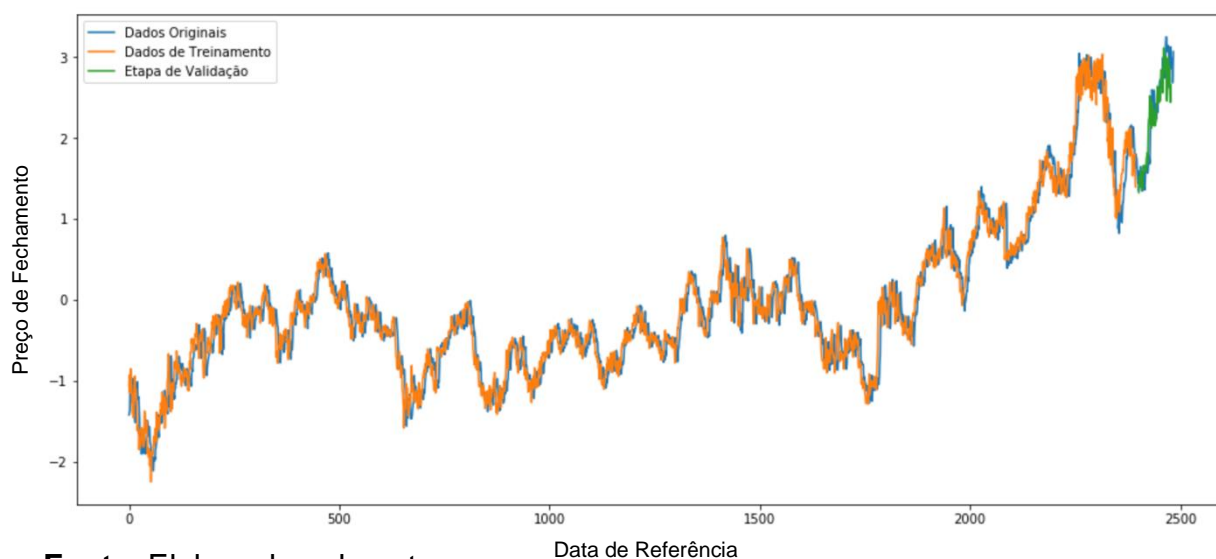
**Tabela 11** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo ITUB4.

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
PIB Real	LSTM considerando somente variáveis técnicas.	D-1	0,43452
		D-15	0,43448
		D-30	0,43339
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0,29043
		D-15	0,29255
		D-30	0,29042
	LSTM considerando todas as variáveis.	D-1	0,42299
		D-15	0,42556
		D-30	0,42234

**Fonte:** Elaborado pelo autor



**Figura 18** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo ITUB4



**Fonte:** Elaborado pelo autor

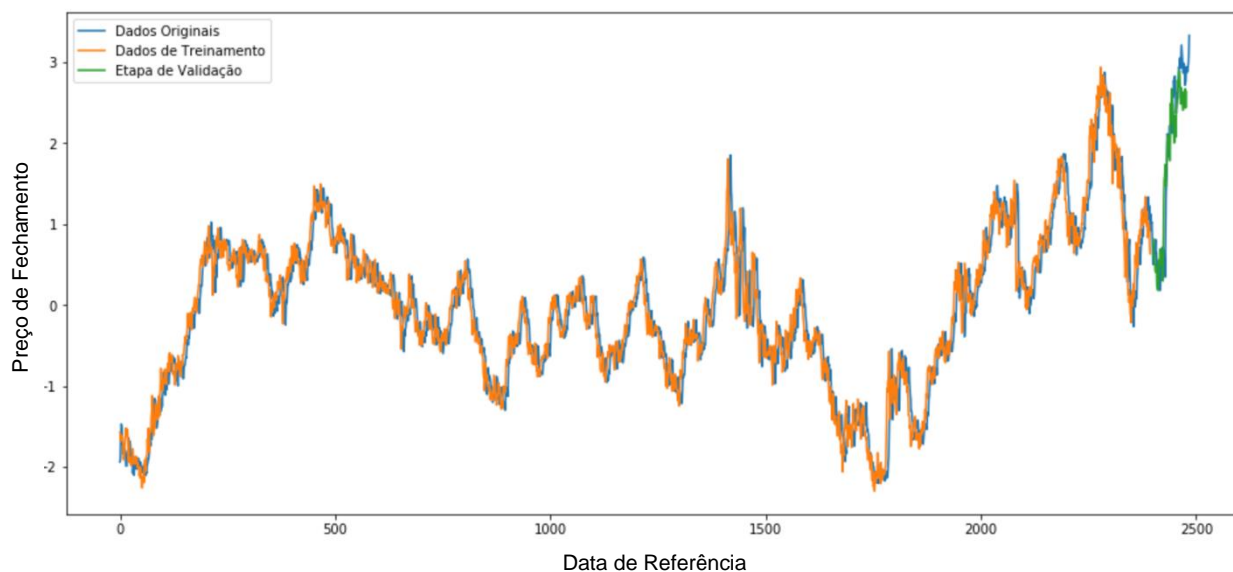
A tabela 12 mostra os resultados para o ativo BBDC4, o gráfico correspondente pode ser visualizado na figura 19.

**Tabela 12** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BBDC4.

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
Índice de Commodities Energéticas	LSTM considerando somente variáveis técnicas.	D-1	0,47298
		D-15	0,47334
		D-30	0,47221
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0,31122
		D-15	0,31222
		D-30	0,31447
	LSTM considerando todas as variáveis.	D-1	0,59131
		D-15	0,59113
		D-30	0,59001

**Fonte:** Elaborado pelo autor

**Figura 19** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BBDC4



**Fonte:** Elaborado pelo autor

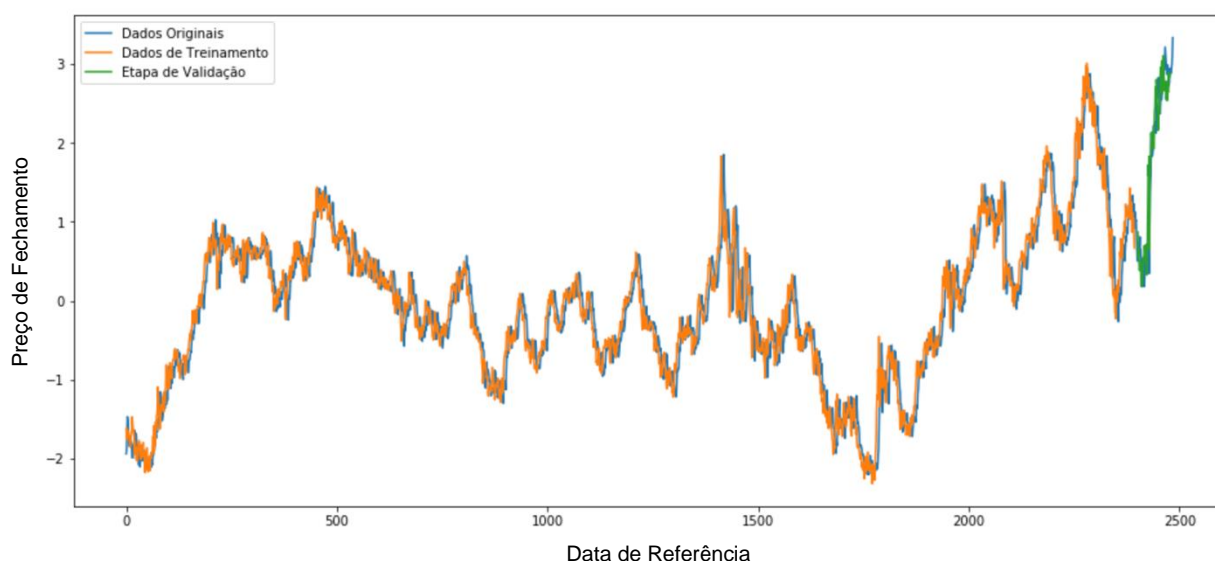
A tabela 13 mostra os resultados para o ativo BBSA3, o gráfico correspondente pode ser visualizado na figura 20.

**Tabela 13** – Demonstração do resultado da rede neural LSTM aplicada para previsão do preço de fechamento – Ativo BBSA3

Variável Macroeconômica com maior correlação com a variável resposta	Configuração Testada	Janela de Tempo	RMSE
Fechamento IBOVESPA	LSTM considerando somente variáveis técnicas.	D-1	0,22014
		D-15	0,22998
		D-30	0,22664
	LSTM considerando variáveis técnicas e a variável macroeconômica com maior correlação.	D-1	0,55773
		D-15	0,55224
		D-30	0,55773
	LSTM considerando todas as variáveis.	D-1	0,54155
		D-15	0,54113
		D-30	0,54994

**Fonte:** Elaborado pelo autor

**Figura 20** – Representação do melhor resultado da previsão do preço de fechamento da rede neural LSTM para o ativo BBSA3



**Fonte:** Elaborado pelo autor

Percebe-se pelos gráficos analisados que a rede neural aplicada foi capaz de prever a variável “Preço de Fechamento” de maneira satisfatória. Após análise individual do desempenho de cada modelo é possível observar a influência positiva no desempenho da rede neural LSTM para os ativos ITUB4 e BBDC4. Conclui-se então que o uso de redes LSTM é adequado para séries temporais financeiras e que variáveis macroeconômicas podem ser utilizadas como características para capturar o efeito momento do preço de fechamento de ações no mercado brasileiro. A tabela 14 mostra o comparativo com os melhores resultados para cada ativo.

**Tabela 14** – Melhores resultados consolidados para os papéis avaliados

Ativo	Variável Macroeconômica com maior correlação	Janela de Tempo	Utilizou variáveis Macroeconômicas?	Melhor RMSE
BOVA11	PIB Real	D-1	Não	0,24523
VALE3	PIB Real	D-30	Não	0,14224
PETR4	Índice Nacional de Expectativas do Consumidor	D-1	Não	0,12245
ITUB4	Fechamento IBOVESPA	D-30	Sim	0,29042
BBDC4	Índice de Commodities Energéticas	D-1	Sim	0,31122

<b>Ativo</b>	<b>Variável Macroeconômica com maior correlação</b>	<b>Janela de Tempo</b>	<b>Utilizou variáveis Macroeconômicas?</b>	<b>Melhor RMSE</b>
BBSA3	Fechamento IBOVESPA	D-1	Não	0,22014

**Fonte:** Elaborado pelo autor

## 6 CONCLUSÕES E RECOMENDAÇÕES

Os resultados dos experimentos executados mostram que redes neurais podem servir para a previsão do preço de fechamento um ativo no mercado de ações e fundamentar uma estratégia de trading desde que os atributos de arquitetura, topologia, amostras de treinamento e variáveis apropriadas. Apesar do desempenho satisfatório o sistema proposto neste estudo pode ser melhorado através da aplicação de outros tipos de ativos, teste com outras variáveis e também com outras janelas de previsão. Os modelos desenvolvidos neste trabalho, baseados em redes neurais mostraram se adequados a previsão das séries temporais propostas, o período de estudo das amostras foi de 03 de dezembro de 2008 a 31 de dezembro de 2018.

Os objetivos definidos no início desta pesquisa foram atendidos. Foi desenvolvido um tipo de rede neural adequado para este tipo de problema classificado como multivariado, onde existe mais de uma observação para cada dado no tempo  $t$ . O algoritmo aplicado apresentou um resultado satisfatório, tanto em termos de RMSE. Desta forma, fica demonstrado a possibilidade da aplicação de redes neurais para a previsão de séries temporais utilizando variáveis macroeconômicas.

Pela evidência apresentada é possível explicar a previsibilidade dos ativos pela pouca oscilação dos dados no período analisado. Infere-se que uma situação de baixa volatilidade em uma série temporal torna a rede neural mais capaz de encontrar padrões entre dados e conferir um grau de previsibilidade não-aleatória. Cabe também a ressalva de que esta pouca volatilidade pode ser explicada devido a políticas macroeconômicas implantadas anteriormente bem como um interesse maior do investidor em ativos comercializados na bolsa de valores. De forma a explorar mais detalhadamente o impacto de variáveis macroeconômicas sugere-se a alteração do algoritmo desenvolvido de forma a realizar previsões em outras janelas temporais, por exemplo, dados intradiários.

Neste estudo também foi possível identificar que as variáveis afetam o preço dos ativos de diferentes maneiras. Portanto, outras técnicas de seleção de parâmetros e combinação de variáveis podem ser estudadas em futuras pesquisas com o objetivo de aumentar a acurácia e encontrar as melhores variáveis que podem representar o efeito momento do preço de fechamento no mercado brasileiro.

Também seria interessante avaliar a correlação e o impacto variáveis não relacionadas ao contexto macroeconômico ou técnico, como por exemplo dados oriundos de análise de sentimento de mercado, menção do nome da empresa em jornais especializados e redes sociais que talvez possam melhorar os resultados apresentados ou causar o efeito oposto.

## REFERÊNCIAS

- ALTINBAS H., BISKIN T. O., **Selecting Macroeconomic Influencers on Stock Markets by Using Feature Selection Algorithms**, Procedia Economics and Finance, Volume 30, 2015.
- ASSAF NETO, A., **Mercado Financeiro**, Atlas, 2001.
- AZOFF E. M., **Neural network time series forecasting of financial markets**. Chichester: John Wiley and Sons, 1994.
- ARÉVALO, R., GARCÍA, J., GUIJARRO, F., PERIS, A., **A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting**. Expert Systems with Applications, 81, 177–192. 2017.
- BETMANN, J. L., SAULT S. J., SCHULTZ E. L., **Fundamental and technical analysis: substitutes or complements?** Accounting and Finance 49, 21–36. 2009.
- CERVELLO-ROYO, R., GUIJARRO F., MICHNIUK K., **Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data**. Expert Systems with Applications, 42(14), 5963–5975. 2015.
- CHEN Y.J., CHEN Y.M., LU. C.L., **Enhancement of stock market forecasting using an improved fundamental analysis-based approach**. Soft Comput (2017) 21: 3735.
- COMISSÃO DE VALORES MOBILIÁRIOS (CVM). **O mercado de valores mobiliários brasileiro**. 3. ed. Rio de Janeiro: Comissão de Valores Mobiliários, 2014.
- CORTEZ, P., RIO M., ROCHA M., SOUSA, P. **Topology aware internet traffic forecasting using neural networks**. International Joint Conference on Neural Networks, Vancouver, 2007.
- COUGHLAN P., COUGHLAN D., **Action Research for Operations Management**. International Journal of Operations & Production Management, v. 22, n. 2, p. 220-240, 2002.
- FAMA, E. F., **The behavior of stock-market prices**. The Journal of Business, 38(1),

34-105.1964.

FAMA, E. F., **Random walks in stock market prices**. Financial Analysis Journal, p. 55–59, 1965.

FAUSETT, L., **Fundamentals of Neural Networks – Architecture, Algorithms, and Applications**. Prentice Hall International, Inc., 1994.

FILHO M. A., ISHIKAWA S. **Mercado Financeiro e de Capitais**. Ed. Atlas; São Paulo; 2000.

FONSECA C. N L., **Análise de desempenho dos fundos multimercados – São Paulo**, 178p. 2012.

FORTUNA E., **Mercado financeiro: Produtos e serviços**. 16. Ed. Rio de Janeiro: Qualitymark, 2005. (2015)

GARCIA, M. E., VASCONCELLOS, M. A. S., **Fundamentos de economia**. São Paulo: Saraiva, 2002.

GAVIN P. H., **The Levenberg-Marquadt method for nonlinear least squares curve-fitting problems**. Dept. Civil and Environmental Engineering Duke Univ. 2011.

GOONTILAKE S., TRELEAVEN P., **Intelligent Systems for Finance and Business**. Chichester, Wiley. 1995.

GRAHAM, B., DODD D. **Security Analysis: The Classic 1934 Edition** (McGraw- Hill, New York). 1934.

GREFF K., SRIVASTAVA K. R., KOTUNÍK J., STEUNEBRINK R. B., SCHMIDHUBER J. **LSTM: A Search Space Odyssey**, IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232, Oct. 2017.

GUNDUZ H., YASLAN Y., CATATELPE Z., **Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations**, Knowledge-Based Systems, Volume 137, 2017.

GURNEY K. (1997). **An introduction to neural networks**, London: UCL Press.



HAYKIN S. **Redes Neurais: princípios e prática**, 2 ed., Bookman, Porto Alegre, RS. 2001.

HAGAN T. M., MENHAJ M., **Training feedforward networks with the Marquardt algorithm**, IEEE Transactions on Neural Networks, 5(6), 989–993, 1994.

HEBB, D. O., **The Organization of Behaviour: A Neuropsychological Theory**, Wiley, New York. 1949.

HINTON G. E., SEJNOWSKI T. J., **Optimal perceptual inference**. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Washington DC.1983.

HOCHREITER S., SCHMIDTHUBER J., **Long short-term memory** (Tech. Rep. No. FKI-207-95). Fakultät für Informatik, Technische Universität München. 1997.

JEGADESH N., TITMAN S., **Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency**. The Journal of Finance, vol. 48, no. 1, 1993.

KAASTRA I., BOYD M., **Designing a neural network for forecasting financial and economic time series**. Neurocomputing: An International Journal, 10(3), 215-236. 1996.

KERR, R. B. **Mercado Financeiro e de Capitais**. São Paulo: Pearson Prentice Hall, 2011.

KINGMA, D., BA, J. **Adam: A Method for Stochastic Optimization**. Proceedings of the 3rd International Conference on Learning Representations ICLR 2015.

KHARE, K., DAREKAR O., GUPTA P., ATTAR V. Z., **Short term stock price prediction using deep learning**, 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2017.

KIRKPATRICK D., DAHLQUIST J., **Technical Analysis: The Complete Resource for Financial Market Technicians**. FT Press, 2nd Edition, 2010.

LACERDA P. D., DRESCH A., PROENCA A., ANTUNES J., **Design Science**

**Research: método de pesquisa para a engenharia de produção.** Gest. Prod., São Carlos, v. 20, n. 4, p. 741-761, 2013 .

LAM, M. **Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis.** Decision Support Systems. 37. 567-581, 2004.

LECUN Y., BOTTOU L., BENGIO P., HAFFNER P., **Gradient-based learning applied to document recognition**, *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

LEITE, H. P., SANVICENTE, A. Z. **Índice Bovespa: um padrão para os investimentos brasileiros.** São Paulo : Atlas, 1994.

LEMOS, F., CARDOSO C. A. **Análise Técnica Clássica**, Editora Saraiva – Brasil: 2010.

LEV B., OHLSON J., **Market-based empirical-research in accounting - A review, interpretation and extension.** Journal of Accounting Research, 20, issue, p. 249-322. 1982.

MARCONI M. A., LAKATOS E. M., **Fundamentos de metodologia científica.** 6ta ed. São Paulo. Atlas. 2008.

MCCULLOCH, W., PITTS, W., **A logical calculus of the ideas immanent in nervous activity**, Bulletin of Mathematical Biophysics 5: 115–133. 1943.

MITCHELL M. T., **Machine learning.** New York: The McGraw-Hill Companies. 1997.

MOORE, M. **Ações: Quais comprar e onde comprar.** Rio de Janeiro, Campus, 2012.

MORETTIN P.A., TOLOI C.M. **Análise de Séries Temporais.** São Paulo, Blucher, 2006.

MURPHY J. J. **Technical Analysis of The Future Markets.** New York Institute of Finance – USA, 1986.

MURPHY J. J. **Charting Made Easy. Trade Secrets** - USA: 2000.

MURPHY J. **Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications.** New York Institute of Finance, 1999.

NAMDARI A., LI S., **Integrating Fundamental and Technical Analysis of Stock Market through Multi-layer Perceptron**, *IEEE Technology and Engineering Management Conference (TEMSCON)*, Evanston, IL, USA, 2018, pp. 1-6.

NAZARIO F. T. R. , E SILVA J. L., SOBREIRO V. A., KIMURA H. **A literature review of technical analysis on stock markets**. *The Quarterly Review of Economics and Finance*, 2017.

OLIVEIRA F. A., NOBRE N.C. e ZÁRATE E. L., **Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index – Case study of PETR4, Petrobras, Brazil**, *Expert Systems with Applications*, Volume 40, Issue 18, 2013.

ROSENBLATT, F. **The perceptron: A probabilistic model for information storage and organization in the brain**, *Psychological Review* 65: 386–408. 1958.

ROYO-C. R., GUIJARRO F., MICHNIUK K. **Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data**. *Expert Systems with Applications*, Volume 42, Issue 14, 2015

SCHALKOF, R. J., **Artificial Neural Networks**. New York, The McGraw-Hill Comp., Inc. 1997.

SCHWAGER J., TURNER S., **Futures: Fundamental Analysis (Fundamental Analysis)**. John Wiley & Sons, 1995.

SELLI M. F., SELEGHIM P. J., **Online identification of horizontal two-phase flow regimes through Gabor transform and neural network processing**. *Heat Transfer Engineering*, v.28, p. 541-548.

SOUSA M. R., VIVIAN A., WOHAR E. M. **Predicting asset returns in the BRICS: The role of macroeconomic and fundamental predictors**, *International Review of Economics & Finance*, Volume 41, 2016.

TKA'C,M., VERNER, R. **Artificial neural networks in business: two decades of research**, *Applied Soft Computing Journal*, 2016.

TRIPPI, R.R., EFRAIM, T., **Neural Network in Finance and Investing**. Chicago, Irwin

Professional Pub. 1996.

WANG, J. L., CHANG, S. H., **Stock market trading rule discovery using pattern recognition and technical analysis**. Expert Systems with Applications. 2007.

WEIGEND, A. S., GERSHENFELD N .A., **Time Series Prediction: Forecasting the Future and Understanding the Past**. Addison-Wesley, 1994.

WHITE, H., **Economic prediction using neural networks: the case of IBM daily stock returns**, IEEE 1988 International Conference on Neural Networks, San Diego, CA, USA, 1988. pp. 451-458 vol.2. 1988.

YU H., WILAMOWSKI M. B., **Levenberg-Marquardt Training Industrial Electronics Handbook**, vol. 5 Intelligent Systems, 2nd Edition, chapter 12, pp. 12-1 to 12-15, CRC Press 2011.. 12-1 to 12. 2010.

ZAHEDI J., ROUNAGHI M. M., **Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange**. Physica A: Statistical Mechanics and its Applications, Volume 438, 2015.

ZHANG, G. P., PATUWO B. E., HU M. Y., **A simulation study of artificial neural networks for nonlinear time-series forecasting**. Atlanta, Computers & Operations Research 28, 2001, p. 381-396.