

Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia Eletrônica

Uso de Redes Neurais Recorrentes Na Predição do Índice do Bovespa

Autor: Felipe Chermont Meneses Pereira
Orientador: Prof. Dr. Gerardo Antonio Idrobo Pizo

Brasília, DF
2023



Felipe Chermont Meneses Pereira

Uso de Redes Neurais Recorrentes Na Predição do Índice do Bovespa

Monografia submetida ao curso de graduação em (Engenharia Eletrônica) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia Eletrônica).

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Prof. Dr. Gerardo Antonio Idrobo Pizo

Brasília, DF

2023

Felipe Chermont Meneses Pereira

Uso de Redes Neurais Recorrentes Na Predição do Índice do Bovespa/ Felipe Chermont Meneses Pereira. – Brasília, DF, 2023-
91 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Gerardo Antonio Idrobo Pizo

Trabalho de Conclusão de Curso –
Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2023.

1. Redes neurais recorrentes. 2. Bolsa de valores. I. Prof. Dr. Gerardo Antonio Idrobo Pizo. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Uso de Redes Neurais Recorrentes Na Predição do Índice do Bovespa

CDU 02:141:005.6

Felipe Chermont Meneses Pereira

Uso de Redes Neurais Recorrentes Na Predição do Índice do Bovespa

Monografia submetida ao curso de graduação em (Engenharia Eletrônica) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia Eletrônica).

Brasília, DF, 24 de Fevereiro de 2023:

**Prof. Dr. Gerardo Antonio Idrobo
Pizo**
Orientador

**Prof. Dr. Marcelino Monteiro de
Andrade**
Convidado 1

Prof. Dr. Renato Coral Sampaio
Convidado 2

Brasília, DF
2023

Agradecimentos

A minha mãe, que me proporcionou a melhor educação possível e me incentivou a estudar da melhor forma que podia.

Agradeço aos meus amigos e colegas que fizeram parte de toda minha jornada universitária.

Agradeço a minha belíssima esposa Leticia que sempre me apoiou e ajudou em todos os sentidos da vida.

Agradeço ao meu orientador, que sempre esteve disponível para me ajudar em todos os passos necessários neste trabalho, me incentivando e orientando para entregar o melhor possível.

Muito obrigado a todos.

Resumo

O mercado de ações desempenha um papel importante no desenvolvimento econômico e o Ibovespa é o principal indicador de desempenho das ações negociadas no Brasil. Porém, a predição de quaisquer tipo de ações ainda é muito difícil por conta de sua sensibilidade à política, economia e fatores sociais. Por estes motivos, tanto no mundo acadêmico quanto corporativo, a predição de séries temporais financeiras vem sendo cada vez mais explorada. Uma das técnicas mais utilizadas para esse tipo de predição é a aplicação de redes neurais artificiais.

Este trabalho foca em uma das classes de redes neurais artificiais, as redes neurais recorrentes, onde pela pesquisa bibliográfica realizada, apresenta os melhores resultados em relação as principais métricas de avaliação observadas. É realizado duas arquiteturas desta área: o modelo GRU (*Gated Recurrent Units*) e a LSTM (*Long Short Term Memory*). Onde procura-se realizar uma predição em diferentes horizontes de tempo da ação BOVA11, que replica o índice do BOVESPA, realizando um comparativo entre ambas arquiteturas, procurando entender como pode ser realizado o uso desse tipo de redes neurais artificiais na predição de séries temporais financeiras.

Comparando ambas redes neurais recorrentes, este trabalho conclui que o modelo GRU apresenta melhores resultados em relação ao modelo LSTM pelas arquiteturas desenvolvidas analisando as métricas MAE, MAPE e RMSE quando observado os horizontes de tempo de 5, 30 e 90 dias, enquanto a modelagem LSTM apresentou melhores resultados quando observado um horizonte de 1 dia previsionado. Conclui-se também que ambas arquiteturas não foram capazes de efetivamente coletar as nuances e altas variações do índice da bolsa de valores brasileira, mantendo suas previsões de forma conservadora e variação relativamente linear.

Palavras-chaves: Predição de ações. Ibovespa. Redes neurais artificiais. Redes neurais recorrentes. Séries temporais financeiras.

Abstract

The stock market plays an important role in economic development and the Ibovespa is the main performance indicator for stocks traded in Brazil. However, the prediction of any type of action is still very difficult, due to its sensitivity to political, economic and social factors. For this reason, both in the academic and corporate worlds, the prediction of financial time series has been increasingly explored. One of the most used techniques for this type of prediction is the application of artificial neural networks.

This work focuses on one of the classes of artificial neural networks, the recurrent neural networks, where, based on the bibliographical research carried out, it presents the best results in relation to the main evaluation metrics observed. Two architectures are developed in this area: the model in GRU (*Gated Recurrent Units*) and the LSTM (*Long Short Term Memory*). Where an attempt is made to make a prediction in different time horizons of the BOVA11 action, which replicates the BOVESPA index, making a comparison between both architectures, trying to understand how this type of artificial neural networks can be used in the prediction of financial time series .

Comparing both recurrent neural networks, this work concludes that the GRU model presents better results in relation to the LSTM model by the architectures developed analyzing the MAE, MAPE and RMSE metrics when observing the time horizons of 5, 30 and 90 days, while LSTM presented better results observing a forecasted 1-day horizon. It is also concluded that both architectures were not able to collect the nuances and exacerbated variations of the Brazilian stock exchange index, keeping its dispositions in a conservative way and relatively linear variation.

Key-words: Stock prediction. Ibovespa. Artificial neural networks. Recurrent neural networks. Financial time series.

Lista de ilustrações

Figura 1 – Neurônio Artificial (FLAUZINO, 2010)	20
Figura 2 – Exemplo de uma rede neural convolucional (VARGAS; CARVALHO; VASCONCELOS, 2017)	21
Figura 3 – Rede neural Recorrente (COLAH, 2015)	22
Figura 4 – Rede Neural Recorrente Expandida (COLAH, 2015)	22
Figura 5 – Exemplo Long Short-Term Memory (COLAH, 2015)	23
Figura 6 – Estrutura de rede GRU (CALVETTE, 2019)	24
Figura 7 – Fluxograma do modelo GA-LSTM (CHUNG; SHIN, 2018)	29
Figura 8 – Repositório da Monografia no GitHub	34
Figura 9 – Listagem de Tarefas com o ZenHub	34
Figura 10 – Diagrama Prisma	42
Figura 11 – Fluxo de Desenvolvimento do modelo de predição	43
Figura 12 – Uma rede de <i>backpropagation</i> totalmente conectada (WYTHOFF, 1993)	44
Figura 13 – Fluxograma de Atividades da Monografia	46
Figura 14 – Exemplo do conjunto de dados (BOVA11) extraídas do site (YAHOO,) , mostrando as 10 primeiras linhas da base de dados	50
Figura 15 – Preço no Fechamento do Pregão do BOVA11 de 2009 a Fevereiro de 2023	52
Figura 16 – Demonstrativo do efeito da média móvel na série temporal	53
Figura 17 – Demonstrativo das bandas de bollinger aplicadas ao longo do tempo da ação do BOVA11	54
Figura 18 – Variação do fechamento de mercado em relação ao dia anterior	55
Figura 19 – Resultados obtidos no teste da predição com a rede LSTM nas ações do Ibovespa	61
Figura 20 – Resultados obtidos no teste da predição com a rede GRU nas ações do Ibovespa	62
Figura 21 – Comparativo entre as arquiteturas para 1 dia previsionado	64
Figura 22 – Comparativo entre as arquiteturas para 5 dia previsionado	65
Figura 23 – Comparativo entre as arquiteturas para 30 dia previsionado	65
Figura 24 – Comparativo entre as arquiteturas para 90 dia previsionado	65
Figura 25 – Comparativo da variação entre dias para 5 dias previsionados	66
Figura 26 – Comparativo da variação entre dias para 30 dias previsionados	66
Figura 27 – Comparativo da variação entre dias para 90 dias previsionados	67

Lista de tabelas

Tabela 1 – Resultados relevantes analisados na pesquisa bibliográfica	31
Tabela 2 – Resumo de Classificações da Pesquisa Científica	40
Tabela 3 – Elaboração da String de Busca	40
Tabela 4 – Cronograma TCC 1	48
Tabela 5 – Cronograma TCC 2	48
Tabela 6 – Tunagem dos hyperparâmetros da rede GRU	58
Tabela 7 – Tunagem dos hyperparâmetros da rede LSTM	59
Tabela 8 – Especificações de Hardware	59
Tabela 9 – Comparativo das métricas de erro coletadas dos testes	62
Tabela 10 – Comparativo dos resultados analisados na pesquisa bibliográfica e do desenvolvido pelo Autor	63
Tabela 11 – Comparativo das métricas de erro entre os modelos para 1, 5, 30 e 90 dias previsionados	64
Tabela 12 – <i>Status</i> das atividades referentes ao TCC 1	69
Tabela 13 – <i>Status</i> das atividades referentes ao TCC 2	69

Lista de abreviaturas e siglas

API	<i>Application Programmimg Interface</i> - Interface de programação de aplicativos
BOVESPA	Bolsa de Valores Brasileira
GRU	<i>Gated Recurrent Units</i> - Unidades recorrentes fechadas
IBGE	Instituto Brasileiro de Geografia e Estatística
IBOVESPA	Índice da Bolsa de Valores do Brasil
KNN	<i>k-nearest neighbors</i> - k-vizinhos mais próximos
KOSPI	<i>Korea Composite Stock Price Index</i> - Bolsa de Valores da Korea
LSTM	<i>Long Short-Term Memory</i> - Memória de Curto Prazo Longa
MAE	<i>Mean absolute Error</i> - Erro Médio Absoluto
MAPE	<i>Mean Absolute Percentage Error</i> - Erro Percentual Absoluto Médio
MSE	<i>Mean Squared Error</i> - Erro Quadrático Médio
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i> - Itens de relatório preferidos para revisões sistemáticas e meta-análises
ReLU	<i>Rectified Linear Unit</i> - Unidade Linear Retificada
RMSE	<i>Root Mean Square Error</i> - Erro da Raiz Média Quadrada
RNA	Rede Neural Artificial
RNN	<i>Recurrent neural network</i> - Rede Neural Recorrente
ROC	<i>receiver operating characteristic</i>
SVC	<i>Support Vector Machines</i> - Máquinas de vetor de suporte
S&P500	<i>Standard and Poor's 500</i>
Tanh	Tangente hiperbólica
TCC	Trabalho de Conclusão de Curso

Lista de símbolos

x_i	sinais de entrada
w_i	pesos sinápticos
Σ	Somatório
θ	Theta
σ	Sigma

Sumário

1	INTRODUÇÃO	14
1.1	Contextualização	14
1.2	Justificativa	15
1.3	Questão de Pesquisa	15
1.4	Objetivos	15
1.4.1	Objetivo Geral	15
1.4.2	Objetivos Específicos	15
1.5	Organização do Trabalho	16
2	REFERENCIAL TEÓRICO	17
2.1	Bolsa de Valores	17
2.1.1	Bolsa de Valores no Brasil	17
2.1.2	Ações	18
2.1.3	Índice Bovespa	18
2.2	Métodos de Deep Learning	19
2.2.1	Redes Neurais Artificiais	19
2.2.2	Redes Neurais Convolucionais	21
2.2.3	Redes Neurais Recorrentes	21
2.2.3.1	Long Short-Term Memory	22
2.2.3.2	Gated Recurrent Units (GRU)	24
2.3	Métricas de Avaliação	25
2.3.1	Erro Médio Absoluto	25
2.3.2	Erro Percentual Absoluto Médio	25
2.3.3	Erro Quadrático Médio	26
2.3.4	Erro Relativo da Raiz Média Quadrada	26
2.3.5	Acurácia	27
2.3.6	Precisão	27
2.3.7	Retorno	28
2.3.8	F1-Scores	28
2.4	Trabalhos Correlatos	28
2.5	Considerações Finais	32
3	REFERENCIAL TECNOLÓGICO	33
3.1	Suporte e Gestão	33
3.1.1	Git 2.37.2	33
3.1.2	GitHub	33

3.1.3	ZenHub	34
3.2	Desenvolvimento Técnico	35
3.2.1	Jupyter Notebook	35
3.2.2	Python 3.10.4	35
3.2.2.1	Pandas 1.4.3	35
3.2.2.2	Numpy 1.23.0	36
3.2.2.3	Matplotlib 3.5.2	36
3.2.2.4	Keras e TensorFlow 2	36
3.3	Considerações Finais	36
4	METODOLOGIA	38
4.1	Pesquisa Científica	38
4.1.1	Abordagem	38
4.1.2	Natureza	39
4.1.3	Objetivos	39
4.1.4	Procedimento	39
4.2	Metodologia de Pesquisa Investigativa	40
4.2.1	<i>String</i> de Busca	40
4.2.2	Guia em diagramação Prisma	41
4.3	Metodologia de Desenvolvimento	42
4.4	Metodologia de Análise de Resultados	44
4.5	Fluxo de Atividades	45
4.6	Cronograma	47
4.7	Considerações Finais	49
5	PROPOSTA	50
5.1	Coleta de Dados	50
5.2	Pré-Processamento	51
5.2.1	Análise Exploratória	51
5.2.2	Engenharia de Variáveis	52
5.3	Modelagem	56
5.3.1	Hiperparâmetros	57
5.3.2	Hardware	59
5.4	Considerações Finais	60
6	RESULTADOS	61
7	CONSIDERAÇÕES FINAIS	68
7.1	Status do Trabalho	68
7.1.1	Trabalhos Futuros	70

	REFERÊNCIAS	71
.1	Slides de apresentação do trabalho final	74

1 Introdução

Este capítulo tem a finalidade de apresentar uma [Contextualização](#) acerca dos assuntos tratados neste trabalho, analisando os comportamentos de uma rede neural recursiva aplicada na predição do índice da bolsa de valores Brasileira. A partir dessa contextualização estabelece-se o âmbito da aplicação do projeto, levando a [Questão de Pesquisa](#) do estudo, permitindo compreender a [Justificativa](#) e os principais [Objetivos](#) desse trabalho. Conclui-se o capítulo apresentando a [Organização do Trabalho](#), com os demais capítulos que constitui essa monografia.

1.1 Contextualização

A previsão de ações concentra-se em estimar o movimento futuro do preço de uma ação, que é geralmente percebida como uma tarefa desafiadora devido ao ruído e características voláteis. A variação dinâmica dos preços do mercado de ações e seu comportamento caótico aumentaram o problema de previsão de preços, onde o domínio não linear, dinâmico e complicado conhecimento inerente ao mercado de ações elevou a dificuldade dos investidores a tomar decisões de investimento rápidas ([CHOPRA; SHARMA, 2021](#)). As previsões tem sido um tópico de pesquisa predominante a muitos anos, quando um sistema pode selecionar de forma confiável as tendências do mercado o sistema pode gerar maior renda para seus proprietários.

Vários sistemas inteligentes e técnicas de inteligência artificial foram desenvolvidas ao longo dos anos para auxiliar as decisões dos investidores ao comprar uma ação. Dentre as técnicas mais utilizadas, as redes neurais artificiais (RNAs) são amplamente populares, devido a capacidade de analisar relações não lineares complexas. A utilização de técnicas de representação do conhecimento para prever preços de ativos tem demonstrado resultados positivos, em que as redes neurais artificiais podem tratar o tema de maneira mais simples que os métodos não lineares ([MARTINS; METTE; MACEDO, 2008](#)).

Neste contexto, este trabalho objetiva a predição do IBOVESPA, sendo este o principal indicador de desempenho das ações negociadas na B3, bolsa de valores brasileira, utilizando a estrutura de rede neural recorrente, com arquiteturas de GRU (*Gated Recurrent Units*) e *Long Short-Term Memory* (LSTM), trazendo um comparativo entre ambas.

1.2 Justificativa

Por conta da volatilidade inerente do mercado de ações, o investidor pode acabar tomando decisões de compra e venda equivocadas por falta de conhecimento do mercado ou especulações, levando o investidor a ter prejuízo.

O trabalho do (NABIPOUR et al., 2020) realiza uma comparação entre nove modelos de *Machine Learning* (Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naive Bayes, KNN, Logistic Regression e RNA) e dois modelos de *Deep Learning* (RNN e LSTM) na predição de ações do mercado de petróleo, finanças gerais, minerais metálicos e não metálicos, em que através das métricas de acurácia, F1-score e ROC, conclui-se que os modelos de RNN e LSTM obtiveram resultados melhores. Porém não encontra-se muitos estudos mais aprofundados no uso de modelos de *Deep Learning* na predição do IBOVESPA.

Esse trabalho tem como principal finalidade a predição com menor erro possível da série temporal em diferentes horizontes de tempo do IBOVESPA, bem como o objetivo de conhecer as tendências do mercado de ações brasileiro, através do índice mais representativo do Brasil, visando maior confiabilidade nas ações de compra e venda, e aprofundar o conhecimento na utilização de redes neurais recorrentes para predição de séries temporais.

1.3 Questão de Pesquisa

Em consideração ao apresentado, esse trabalho de monografia visa responder a seguinte questão de pesquisa: **É possível utilizar redes neurais recorrentes para prever o índice da bolsa de valores brasileira?**

1.4 Objetivos

1.4.1 Objetivo Geral

Utilizando dados históricos do IBOVESPA, pretende-se com este trabalho cumprir o objetivo geral de estudar a viabilidade do uso de redes neurais recorrentes para prever o índice da bolsa de valores brasileira, analisando as principais métricas de avaliação e realizando um comparativo entre as arquiteturas de *Long Short-Term Memory* (LSTM) e GRU (*Gated Recurrent Units*) em diferentes horizontes de tempo.

1.4.2 Objetivos Específicos

No intuito de cumprir o objetivo geral proposto, este trabalho pretende:

- Avaliar os dados históricos de volume negociado, valor na abertura, valor de fechamento, máxima no dia e mínima no dia, do IBOVESPA através de ferramentas estatísticas;
- Verificar o efeito das técnicas de pré-processamento;
- Realizar um estudo dos hiperparâmetros, suas influências na rede e a melhor combinação possível para cada modelo;
- Avaliar a qualidade das previsões num espaço de tempo de 1, 5, 30 e 90 dias à frente do último valor usado para treinamento;
- Comparar as técnicas e resultados obtidos com a literatura analisada na pesquisa bibliográfica.

1.5 Organização do Trabalho

Esta monografia está estruturada conforme os capítulos:

- Capítulo 1 - **Introdução**: introduz o tema de pesquisa, com a contextualização do cenário analisado, a justificativa, a questão de pesquisa, os objetivos gerais e específicos, e a composição da organização do trabalho;
- Capítulo 2 - **Referencial Teórico**: expõe a base teórica para desenvolvimento do trabalho, detalhando os principais tópicos da Bolsa de Valores brasileira, o índice BOVESPA que será trabalhado, e os principais métodos de *Deep Learning*, com foco nas Redes Neurais Recorrentes, que serão trabalhadas ao longo do estudo;
- Capítulo 3 - **Referencial Tecnológico**: mostra as tecnologias e aplicações utilizadas no desenvolvimento do trabalho
- Capítulo 4 - **Metodologia**: exhibe a metodologia que guia a pesquisa, o desenvolvimento e a análises do estudo;
- Capítulo 5 - **Proposta**: indica a proposta deste trabalho, definindo o desenvolvimento realizado;
- Capítulo 6 - **Resultados** Apresenta os resultados coletados pela proposta definida e uma discussão dos efeitos da aplicação.
- Capítulo 7 - **Considerações Finais** apresenta as considerações e conclusões de todo o desenvolvimento obtidos durante a aplicação do estudo.

2 Referencial Teórico

Com o objetivo de integrar-se nos principais conceitos de interesse desse trabalho, este capítulo aborda a definição e conceito sobre: [Bolsa de Valores](#), [Métodos de Deep Learning](#) e [Métricas de Avaliação](#). Com o intuito de conferir os fundamentos para orientação da pesquisa, assim como a elaboração da modelagem pretendida, direcionando-se pela literatura especializada, em [Trabalhos Correlatos](#). Concluindo com a apresentação das [Considerações Finais](#) do capítulo.

2.1 Bolsa de Valores

De forma geral, uma bolsa de valores é um tipo de mercado onde são efetuadas transações de compra e venda de produtos agrícolas e matérias primas ou valores mobiliários. Geralmente representados por títulos de empresas privadas e instituições governamentais, incluem os debêntures e as ações. Entende-se por mercado, o local onde são realizadas as transações, as pessoas que as realizam e o conjunto destas transações. (GOMES, 2007)

2.1.1 Bolsa de Valores no Brasil

Dentre os atos legais acerca das bolsas de valores no Brasil, pode-se destacar a Resolução nº39, de 20 de outubro de 1966, do Banco Central, que regulamentou a norma que disciplina a constituição, organização e funcionamento destas instituições em todo o país. (GOMES, 2007)

De acordo com essa Resolução: *“As Bolsas de Valores são associações civis, sem finalidades lucrativas, tendo por objeto social manter local adequado ao encontro de seus Membros e à realização, entre eles, de transações de compra e venda de títulos e valores mobiliários, em mercado livre e aberto, especialmente organizado e fiscalizado por seus membros e pelas autoridades monetárias”*(GOMES, 2007)

No Brasil atualmente, as bolsas de valores praticamente negociam apenas ações e fundos imobiliários, que consistem em títulos que dão direito de propriedade sobre parte de uma companhia as pessoas que as adquirirem. Estes títulos, antes de negociados devem ser registrados na bolsa. Que por sua vez, efetua a vistoria nas companhias que optarem por ações em bolsa. (GOMES, 2007)

2.1.2 Ações

Uma ação representa a menor parcela do capital de uma empresa. Quem compra ações de uma companhia adquire também os direitos e os deveres de um sócio. Se for uma companhia aberta registrada na Comissão de Valores Mobiliários (a CVM, órgão que regula e fiscaliza o mercado de capitais brasileiro), as ações podem ser negociadas publicamente na bolsa de valores. (INFOMONEY, 2022)

Uma das principais vantagens de se tornar acionista de uma empresa é poder se beneficiar de parte dos resultados que ela obtiver. Quando uma companhia aberta tem lucro, uma parcela dele é distribuída aos sócios na forma de dividendos, na proporção do número de ações que cada um possui. O acionista pode ganhar também com a possível valorização do preço dos papéis que, além do desempenho financeiro da empresa, depende também das perspectivas para o setor em que ela atua e para a economia em geral. (INFOMONEY, 2022)

Existem dois modelos diferentes que dominam a compra e venda de ações no mercado financeiro: modelo fundamentalista e modelo técnico. O modelo fundamentalista analisa aspectos da empresa em questão e aspectos macroeconômicos. Já o modelo técnico se fundamenta em dados históricos a respeito da empresa, tentando, assim, inferir o futuro a partir de comportamentos passados (AIRES; DAMETTO; CREPALDI, 2017). O histórico de preços de uma determinada ação, portanto, produz uma série temporal não-linear, cuja previsão de comportamento é de fundamental importância para negociações futuras (FLAUZINO, 2010).

2.1.3 Índice Bovespa

O Índice Bovespa, também conhecido como Ibovespa é o principal indicador de desempenho das ações negociadas na B3 e reúne as empresas mais importantes do mercado de capitais brasileiro. Foi criado em 1968 e, ao longo desses 50 anos, consolidou-se como referência para investidores ao redor do mundo. (B3, 2022)

O Índice Bovespa é escolhido em função do elevado volume negociado, da série histórica longa e consistente, de sua utilização para avaliar os produtos de rendas variáveis e da associação com vários papéis futuros, termos e opções (SANVICENTE, 1994). Desta forma, consideramos o índice ibovespa como a melhor forma de representar o mercado brasileiro, adotado por diversos investidores.

Sua composição é formada com ações que obtiveram uma presença de pelo menos 80% das sessões de pregões no período e alcançaram 80% de participação acumulada em termos de números de negócios e volume financeiro nos 12 meses anteriores e também devem possuir uma participação superior a 0,1% do volume total de negócios. Quadri-mestralmente é feita uma reavaliação do mercado, com base nos 12 meses anteriores, para

manter uma representatividade do Índice ao longo do tempo.(SOUZA, 2008)

Atualmente, em Agosto de 2022, o índice Ibovespa é formado por 90 ações, com sua maior posição alocada nas ações: VALE3, PETR4, ITAUB4 e BBDC4, que correspondem respectivamente as empresas, Vale S.A, Petrobrás, Itaú Unibanco e Bradesco, que juntas somam cerca de 30% da composição do índice. Toda a composição do índice está disponível disponível para consulta na (B3, 2022).

2.2 Métodos de Deep Learning

Aprendizagem profunda, mais conhecido como *Deep Learning* do inglês, é um tipo de aprendizado de máquina que treina computadores para realizar tarefas humanas(SAS, 2022).

Deep Learning permite modelos computacionais compostos várias camadas de processamento aprendam e reconheçam padrões afim de representar os dados com diversos níveis de abstração ajustando parâmetros básicos, ao invés da utilização de algoritmos com equações pré-determinadas.

Esses métodos vem ganhando espaço no mercado e sendo utilizando em diversas aplicações como o reconhecimento de fala, reconhecimento visual de objetos, detecção de objetos e muitos outros domínios, como a descoberta de drogas e genômica. (LECUN; HINTON, 2015)

2.2.1 Redes Neurais Artificiais

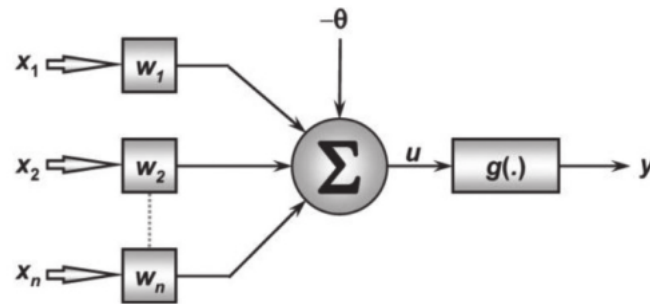
As Redes Neurais Artificiais, conhecidas como RNAs, foram desenvolvidas na década de 1940 por Walter Pitts e McCulloch. O seu desenvolvimento tinha por objetivo o desenvolvimento de um modelo computacional análogo a neurônios biológicos e circuitos eletrônicos, capazes de simular conexões sinápticas pelo uso de resistores variáveis e amplificadores.(JR; COSTA; E.M.M, 2007)

O conceito das RNAs consiste em substituir o desenvolvimento de um computador que de forma inteligente, procura formas de resolver problemas RNAs são formadas por neurônios interligados entre si e sinapses artificiais estruturados com base em modelos biológicos. Desta forma, as RNAs possuem características como capacidade de aprendizagem através de treinamento, habilidade de generalização e tolerância a falhas.(M.A.S; F; MACEDO, 2008)

A interpretação de um neurônio pode ser visualizada conforme a Figura 1, com os x_i sinais de entrada externos e os w_i pesos sinápticos, que ponderam e variam de acordo com a relevância de cada x_i . Logo após, é realizado a soma ponderada dos sinais de entrada pelo combinador linear \sum , onde é subtraído a variável θ conhecida como o limiar

de ativação. Por fim, a saída de cada neurônio é chamada de potencial de ativação u . As saídas dos neurônios artificiais são enviadas para uma função de ativação denominada por $g(\cdot)$, com a função de limitar a saída do neurônio dentro de um intervalo determinado, o qual possa ser assumido pela rede e assim gerar o sinal de saída y . (AIRES; DAMETTO; CREPALDI, 2017)

Figura 1 – Neurônio Artificial (FLAUZINO, 2010)



A representação matemática do relacionamento entre os nós, pesos e vieses é descrita na Equação 2.1. A soma dos pesos de entrada para uma camada passa pela função de ativação não linear para outro nó da próxima camada. O que pode ser interpretado como um vetor.

$$Z = f(x.w + b) = f\left(\sum_{i=1}^n (x_i^T w_i + b)\right) \quad (2.1)$$

onde:

- x_i são os *inputs*;
- w_i são os pesos;
- b é o viés;
- f é a função de ativação;
- Z é a saída.

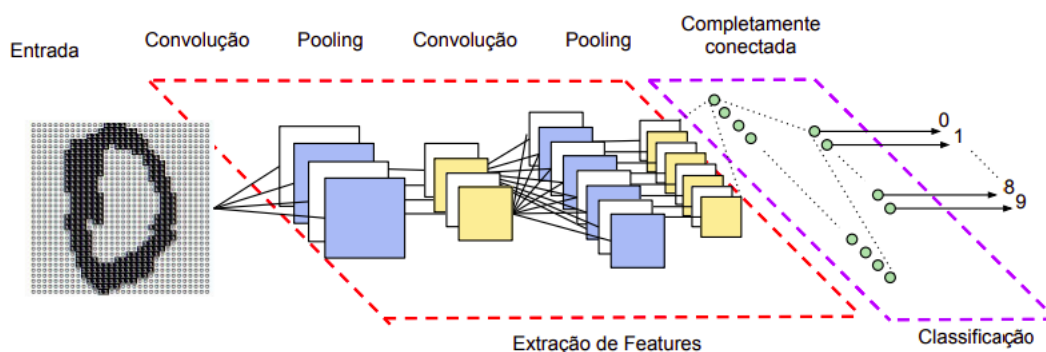
Calculando os pesos e vieses, o processo de treinamento é completado quando: A inicialização dos pesos e vieses de todos os nós deve ser feita de forma aleatória, realizando uma passagem para frente dos atuais pesos e vieses e calculando a saída de cada nó, e modificando os pesos e vieses consequentemente pelo gradiente descendente com passagem para trás, processo conhecido pelo termo em inglês *backpropagation algorithm* (NABIPOUR et al., 2020)

2.2.2 Redes Neurais Convolucionais

Redes Neurais Convolucionais, mais conhecidas como CNN's, originda do termo em inglês (*Convolutional Neural Network*) é muito inspirada no processo biológico de processamentos de dados visuais. Uma CNN é capaz de aplicar filtros de dados visuais, mantendo semelhança de vizinhança entre os pixels da imagem ao longo do processamento da rede, o que permitiu o método ganhar bastante espaço no ambito de visão computacional com aplicações na classificação, detecção e reconhecimento em imagens e vídeos. (VARGAS; CARVALHO; VASCONCELOS, 2017)

Uma CNN é composta pelas seguintes camadas: camada convolucional, camada de agrupamento e camada totalmente conectada. Onde a camada convolucional utiliza filtros sobre a imagem afim de obter uma série de mapas de características para cada filtro. As camadas de agrupamento são responsáveis por reduzir a amostra para os mapas de características e, ao fim, a tarefa de classificação das imagens é realizada pela camada totalmente conectada (SOUZA et al., 2020). Podemos ver um exemplo de uma CNN e suas diversas camadas na Figura 2

Figura 2 – Exemplo de uma rede neural convolucional (VARGAS; CARVALHO; VASCONCELOS, 2017)



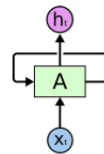
Além da sua vasta aplicação no ambito da visão computacional, as variações de arquiteturas de Redes Neurais Convolucionais veem obtendo bons resultados em outras áreas de aplicação. (KALCHBRENNER; GREFFENSTETTE; BLUNSOM, 2014) utilizaram CNNs Dinâmicas para uma série de tarefas de Processamento Natural de Linguagens como análise de sentimentos e classificação de tipos de pergunta. (WALT et al., 2014) utilizaram o método para classificação binária de frases como positivo ou negativo. (FERREIRA, 2017)

2.2.3 Redes Neurais Recorrentes

As redes neurais recorrentes, do inglês (*recurrent neural networks* - RNN) são um caso especial onde existe realimentação de entradas (A.C, 2006) conforme ilustrado na

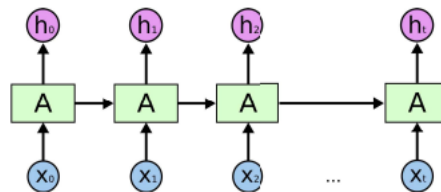
Figura 3. A realimentação faz com que as redes neurais apresentem um comportamento dinâmico temporal. Por apresentarem capacidade de memória, ou seja, habilidade de armazenar informação de dados anteriores, são muito utilizadas em questões relacionadas a séries temporais. Uma rede neural recorrente pode responder a mesma entrada com diferentes saídas em diferentes momentos, dependendo da entrada apresentada anteriormente (PENNA, 2004).

Figura 3 – Rede neural Recorrente (COLAH, 2015)



A figura 3 mostra uma exemplificação de uma RNN em um instante de tempo t e uma entrada x_i produzindo uma saída h_t . Onde a camada A é recorrente. Outro modo de visualizar as redes neurais é expandindo suas recorrências, como demonstrado na figura 4. Desta forma é possível visualizar o fluxo de dados de uma RNN, sendo essa uma expansão da rede mostrada na 3;

Figura 4 – Rede Neural Recorrente Expandida (COLAH, 2015)

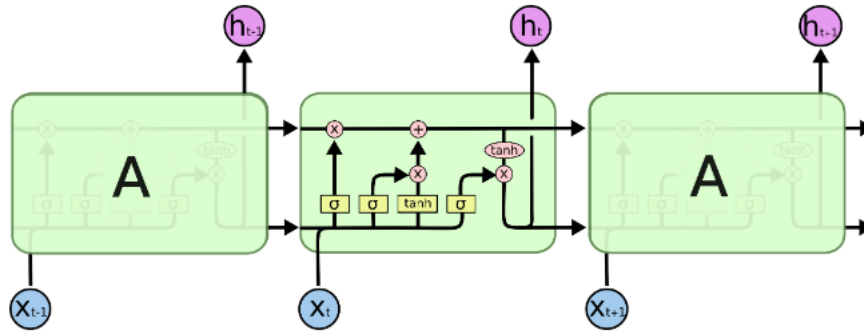


2.2.3.1 Long Short-Term Memory

As redes *Long short-term Memory* (LSTM) são um tipo especial de RNN que permite solucionar problemas de explosão e desvanecimento de gradientes em RNNs, capazes de aprender informações de longo prazo, sendo primeiramente introduzidas por Hochreiter e Schmidhuber (J.HOCHREITER, 1991). LSTMs são bastante utilizadas quando o volume de dados é muito grande e as informações anteriores tornam-se de longo prazo, geralmente apresentando resultados melhores em comparação as RNNs.

A principal diferença entre as RNNs e uma LSTM é que em cada neurônio da LSTM há uma célula de memória. Onde cada neurônio possui três *gates*, sendo eles: *gate* de entrada, *gate* de esquecimento e *gate* de saída (NABIPOUR et al., 2020), conforme ilustrado na Figura 5, onde é utilizado a função sigmoide σ nos *gates* de regulagem e a função *tanh* nos *gates* análogos a função de recompensa.

Figura 5 – Exemplo Long Short-Term Memory (COLAH, 2015)



A primeira operação realizada dentro da LSTM é a remoção das memórias que deveriam ser omitidas da célula, executada pelo *gate* de esquecimento representado na equação 2.2.

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f) \quad (2.2)$$

Em que:

- f_t representa o *gate* de esquecimento;
- X_t se refere *inputs*;
- h_{t-1} se refere ao estado anterior;
- W_f e b_f representam o peso e parâmetro de vetores de viés que precisam ser aprendidos durante o treinamento, respectivamente.

Para definir as informações que será armazenada no estado da célula, é preciso realizar dois processos. Primeiramente a camada chamada de *gate* de entrada decide qual valor será atualizado, logo após, uma camada tangencial cria um vetor \tilde{C}_t e em seguida é combinado o *gate* de entrada com o vetor para realizar a atualização do estado, conforme mostrado na Equação 2.3 e na Equação 2.4.

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \quad (2.3)$$

Onde i_t representa o *gate* de entrada e atualização.

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, X_t] + b_C) \quad (2.4)$$

Em que \tilde{C}_t é a célula de entrada.

é responsável por reiniciar a célula, apagando os períodos passados. A atualização da memória (\tilde{h}_t) realiza uma nova combinação entre a entrada e a memória passada, através da informação passada pelo portão *reset*. Por fim a saída do período atual (h_t) pode ser expressa como uma combinação entre atualização da memória da célula e da informação mantida anteriormente, conforme a equação 2.8.

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (2.8)$$

2.3 Métricas de Avaliação

Algoritmos de inteligência artificial e *Deep Learning* são averiguados em todas as fases do processo, desde a validação à testagem. Diversas linguagens possuem *frameworks* que auxiliam nos cálculos de diversos parâmetros que podem ser úteis na avaliação do algoritmo. É usual interpretações simplistas que exibem exclusivamente situações de classificações binárias, e conforme o aumento da complexidade, são comuns equívocos na classificação. (JUNIOR et al., 2022)

Neste contexto, este tópico visa exemplificar e definir os principais parâmetros de avaliação utilizados nos classificadores de *Machine Learning* e *Deep Learning*, sendo eles: Erro Percentual Absoluto médio, Erro Médio Absoluto, Erro Relativo da Raiz Média Quadrada, Erro Quadrático Médio, Acurácia, Precisão, Recall e F1-Scores.

2.3.1 Erro Médio Absoluto

O Erro Médio Absoluto, do inglês *Mean absolute Error* (MAE) é uma métrica de avaliação bastante usada na avaliação de modelos. O MAE é medido pela média entre dois valores, sendo eles A_t e F_t , representando respectivamente o valor real e o valor da predição, como demonstrado na equação 2.9.

$$MAE = \frac{1}{n} \sum_{t=1}^n | A_t - F_t | \quad (2.9)$$

Onde n representa o número de amostras.

2.3.2 Erro Percentual Absoluto Médio

O Erro Percentual Absoluto Médio, mais conhecido pelo seu termo em inglês *Mean Absolute Percentage Error* (MAPE) é uma medida de erro relativo, que usa valores absolutos para impossibilitar que os erros positivos e negativos se anulem e assim, usa os erros relativos para permitir que haja uma comparação entre a precisão da previsão

nos métodos de séries temporais (TEAM, 2022), sendo sempre apresentada em forma de porcentagem, conforme a Equação 2.10.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left(\left| \frac{A_t - F_t}{A_t} \right| \right) * 100 \quad (2.10)$$

Onde:

- A_t é o valor atual;
- F_t é o valor previsto

Pela fórmula, O valor absoluto é a diferença entre o valor real e previsto dividido pelo valor real. Por fim, o erro é multiplicado por 100 para ficar na forma percentual.

2.3.3 Erro Quadrático Médio

O Erro Quadrático Médio (*Mean Squared Error*, MSE) mensura a qualidade das predições e o valor não negativo, onde quanto mais perto de 0 for a métrica, melhor será a qualidade das predições do modelo. O MSE demonstra o quão amplamente espalhadas as previsões são de uma amostra para outra, e o quão próximo o valor médio previsto está da observação (J et al., 2015). Como pode ser exemplificado pela Equação 2.11

$$MSE = \frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2 \quad (2.11)$$

Onde:

- A_t representa o valor real;
- F_t o valor previsto;
- n o tamanho da amostra

2.3.4 Erro Relativo da Raiz Média Quadrada

O Erro da Raiz Média Quadrada (RMSE) do inglês *Root Mean Square Error* representa o desvio padrão dos erros de previsão no modelo de regressão trabalhado. Os Erros de previsão, mostram a distância entre os valores reais e os valores previstos na modelagem. A métrica é utilizada para indicar como está a concentração dos dados perto do melhor modelo treinado. O Erro Relativo da Raiz Quadrada, ou RRMSE do inglês *Relative Root Mean Square Error* seria uma variação do RMSE, onde o erro quadrático total é normalizado e dividido pelo erro quadrático total previsto (NABIPOUR et al., 2020), como descrito pela Equação 2.12

$$RRMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{A_t - F_t}{A_t} \right)^2} \quad (2.12)$$

Onde:

- A_t representa o valor real;
- F_t o valor previsto;
- n o tamanho da amostra

2.3.5 Acurácia

A acurácia é amplamente utilizada como parâmetro central na avaliação de classificadores na inteligência artificial, principalmente pela sua simplicidade de compreensão. (JUNIOR et al., 2022)

A Acurácia mede o percentual de casos verdadeiros em relação a todos os resultados, podendo ser calculada então como a razão entre o número total de observações certas do modelo e o total de observações previstas, conforme a Equação 2.13.

$$Acuracia = \frac{V_p + V_n}{V_p + F_n + F_p + V_n} \quad (2.13)$$

Onde:

- V_p é o número de amostras com classificação correta positivas;
- V_n o número de amostras com classificação correta negativas;
- F_p número de amostras que o modelo previu como verdadeiro quando o valor real é negativo;
- F_n que é o número de amostras que o modelo previu negativo quando o valor real da classe era positivo.

2.3.6 Precisão

A precisão mede a relação entre as amostras previstas como verdadeiras e positivas e o total previsto. Conforme a Equação 2.14

$$precisão = \frac{V_p}{V_p + F_p} \quad (2.14)$$

Onde:

- V_p são os verdadeiros positivos;
- F_p são os falso positivos

Podemos perceber pela relação que ao elevar a quantidade de falso positivos, a precisão do algoritmo cai. Geralmente a precisão é avaliada quando o peso dos falso positivos é relativo para o problema que se quer solucionar.

2.3.7 Retorno

O retorno, ou como é mais conhecido do inglês *Recall* trata da porcentagem de todos os verdadeiro positivos previstos corretamente. Conforme a Equação 2.15

$$retorno = \frac{V_p}{V_p + F_n} \quad (2.15)$$

Onde:

- V_p representam os verdadeiros positivos;
- F_n representam os falso negativos

Pela relação demonstrada, o peso dos F_n é maior em relação ao F_p no modelo avaliado. Mede a taxa de verdadeiro positivos que o modelo prediz, possibilitando analisar a capacidade do classificador na identificação dos F_n (JUNIOR et al., 2022).

2.3.8 F1-Scores

O F1 score funciona como a média harmônica entre a precisão e o retorno, funciona como uma métrica unificada de ambos, atribuindo um peso maior aos valores menores, em relação a uma média aritmética simples. Como demonstrado na Equação 2.16

$$F1 = \frac{2 * precisão * retorno}{precisão + retorno} \quad (2.16)$$

2.4 Trabalhos Correlatos

A predição do mercado de ações em qualquer ambito tem sido um tópico bastante explorado e pesquisado a muito tempo, embora a maior parte das pesquisas do uso de técnicas de Machine Learning e principalmente Deep learning no mercado financeiro é relativamente recente.

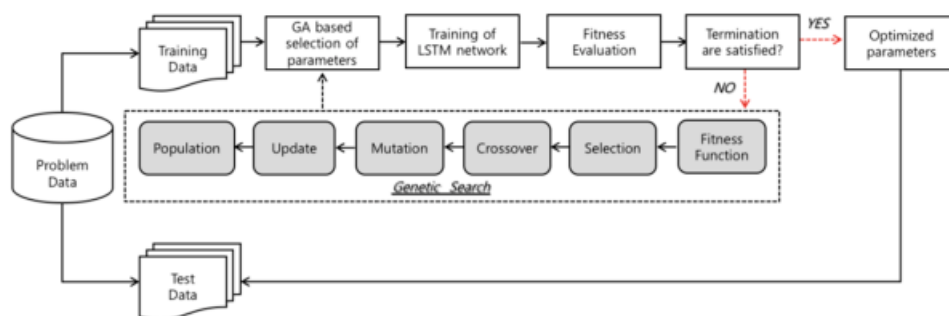
O estudo de (NABIPOUR et al., 2020) utiliza dados históricos de 10 anos dos índices nos grupos de finanças, petróleo, minerais não metálicos e metais básicos da Bolsa

de Valores de Teerã para prever dados de 1, 2, 5, 10, 15, 20 e 30 dias a frente, realizando uma comparação entre diversos tipos de algoritmos de *Machine* e *Deep Learning*, como: *decision tree*, *bagging*, *random forest*, *adaptive boosting (Adaboost)*, *gradient boosting*, and *eXtreme gradient boosting (XGBoost)*, *artificial neural networks (ANN)*, *recurrent neural network (RNN)* e *long short-term memory (LSTM)*. Para mensuração e avaliação dos resultados, foram utilizadas algumas das medidas de erro descritas no capítulo 2.3 sendo as abordadas pelo estudo: MAPE, MAE, RRMSE, MSE. O estudo realizado, chegou ao resultado de que pelos algoritmos e dados utilizados, LSTM apresenta resultados mais acertivos e uma alta capacidade de ajuste do modelo.

O artigo (WANG et al., 2022) avalia a série temporal diária do valor de fechamento de alguns dos principais índices mundiais em termos de volumetria de negociação como o S&P500, Down Jones, Nasdaq e Shanghai Stock Exchange entre as datas de 4 de Janeiro de 2010 a 31 de Dezembro de 2018, observando 2272 dias de transações. O estudo traz uma comparação entre o modelo desenvolvido, chamado de *RC Model* e os métodos mais utilizados de RNN e LSTM, onde o modelo proposto apresenta resultados bastante próximos e em alguns aspectos até mesmo melhor que das arquiteturas convencionais de LSTM e RNN.

O trabalho de (CHUNG; SHIN, 2018) apresenta um estudo utilizando uma abordagem híbrida, integrando uma rede LSTM com um algoritmo genético (GA) na predição do índice de da Bolsa de Valores da Korea (KOSPI), conforme o fluxograma da Figura 7.

Figura 7 – Fluxograma do modelo GA-LSTM (CHUNG; SHIN, 2018)



O estudo contou com dados de Janeiro de 2000 a Dezembro de 2016, contando com 4203 dias de transações, analisando dados de: valor de venda diário, menor preço, maior preço, preço de abertura, preço de fechamento e volume transacionado. Foi analisado as métricas de MSE, MAE e MAPE, onde o modelo teve resultados consideravelmente melhores em relação ao *Benchmark* comparado.

O artigo elaborado por (LAWI; MESRA; AMIR, 2022) realizou um estudo comparando diversas arquiteturas de LSTM e GRU. O artigo analisa dados do dia 03 de Março de 2010 ao dia 17 de Setembro de 2019, totalizando 2403 dados de treinamento das ações: AMZN e GOOGL, BLL, QCOM. Para avaliar os modelos, foram avaliados as métricas

de erro: MAPE, RMSPE e RMSDE, onde foi analisado que ambos modelos tiveram bons resultados em todas as arquiteturas analisadas, mantendo acurácias na faixa de 92,74% a 98,47% para o MAPE percentual, 90,44% a 97,96% para o RMSPE e 66,64% a 90,73% para RMSDE.

Os principais resultados dos estudos citados anteriormente e outros relevantes para composição do trabalho podem ser visualizados na tabela 10 , onde é apresentado o Índice ou ação de estudo, os modelos analisados, as métricas de avaliação, os resultados obtidos e as referências de quem realizou cada pesquisa.

Tabela 1 – Resultados relevantes analisados na pesquisa bibliográfica

Predição	Modelo	Métrica	Resultado	Literatura
Finanças Diversificadas	Decision Tree	MAPE	2,07	(NABIPOUR et al., 2020)
		MAE	35,82	
		RRMSE	0,0396	
		MSE	7984,3	
	Random Forest	MAPE	1,91	
		MAE	31,96	
		RRMSE	0,0288	
		MSE	3962,97	
	Adaboost	MAPE	1,59	
		MAE	26,27	
		RRMSE	0,0248	
		MSE	2867,63	
	RNN	MAPE	1,85	
		MAE	17,20	
		RRMSE	0,0281	
		MSE	635,85	
LSTM	MAPE	0,6		
	MAE	6,7		
	RRMSE	0,0093		
	MSE	148,77		
Korea Stock Price Index (KOSPI)	GA-LSTM	MAPE	0,91	(CHUNG; SHIN, 2018)
		MAE	10,21	
		MSE	181,99	
	Benchmark	MAPE	1,10	
		MAE	11,71	
		MSE	181,99	
AMZN	LSTM	MAPE	0,97	(LAWI; MESRA; AMIR, 2022)
		RRMSE	0,96	
	GRU	MAPE	0,94	
		RRMSE	0,93	
Standard and Poor's 500 (S&P500)	LSTM	MAE	33,5	(LV et al., 2022)
		RMSE	53,4	
	GRU	MAE	43,3	
		RMSE	63,2	
	ARIMA	MAE	34,1	
		RMSE	54,8	
Standard and Poor's 500 (S&P500)	LSTM	MAE	15,66	(WANG et al., 2022)
		RMSE	23,89	
	RC Model	MAE	15,80	
		RMSE	23,25	
	RNN	MAE	15,37	
		RMSE	23,44	

2.5 Considerações Finais

O capítulo apresentou definições e contextualização sobre os principais tópicos abordados nesse trabalho. Inicia-se com a contextualização do mercado de ações e como ela funciona no Brasil, é descrito a motivação para predição do índice do Bovespa e sua composição atual, então, é exemplificado os principais métodos de Deep Learning para análise predição de séries temporais, as métricas de avaliação de modelos de inteligência artificial e por fim, é descrito os principais trabalhos correlatos com o objetivo deste trabalho.

Com a análise dos estudos, percebe-se como as técnicas foram aplicadas para a problemática da predição do valor das ações e os resultados obtidos por cada estudo. É entendido as *features* utilizadas e como foi mensurado os resultados de cada modelo, desta forma, é possível utilizar as melhores ideias dos estudos analisados durante a pesquisa para composição deste trabalho.

3 Referencial Tecnológico

Este capítulo apresenta as principais ferramentas e tecnologias utilizadas no desenvolvimento deste trabalho. A seção [Suporte e Gestão](#) descreve as ferramentas utilizadas no suporte ao armazenamento e controle de gestão do algoritmo e dados do trabalho desenvolvido. Por conseguinte, a seção [Desenvolvimento Técnico](#) confere as principais ferramentas que estão sendo utilizadas no desenvolvimento técnico deste trabalho, visando a implementação do algoritmo de predição. Por fim, conclui-se o capítulo com a seção de [Considerações Finais](#).

3.1 Suporte e Gestão

As ferramentas de suporte e gestão são utilizadas para proporcionar a melhor organização e controle do trabalho. Utiliza-se o [Git 2.37.2](#) e [GitHub](#) para versionamento de código e registro do template utilizado na monografia. Adicionalmente, foram registradas as tarefas no quadro utilizando a ferramenta [ZenHub](#) provinda da metodologia Kanban, para ter o melhor controle das ações e prazos das tarefas realizadas.

3.1.1 Git 2.37.2

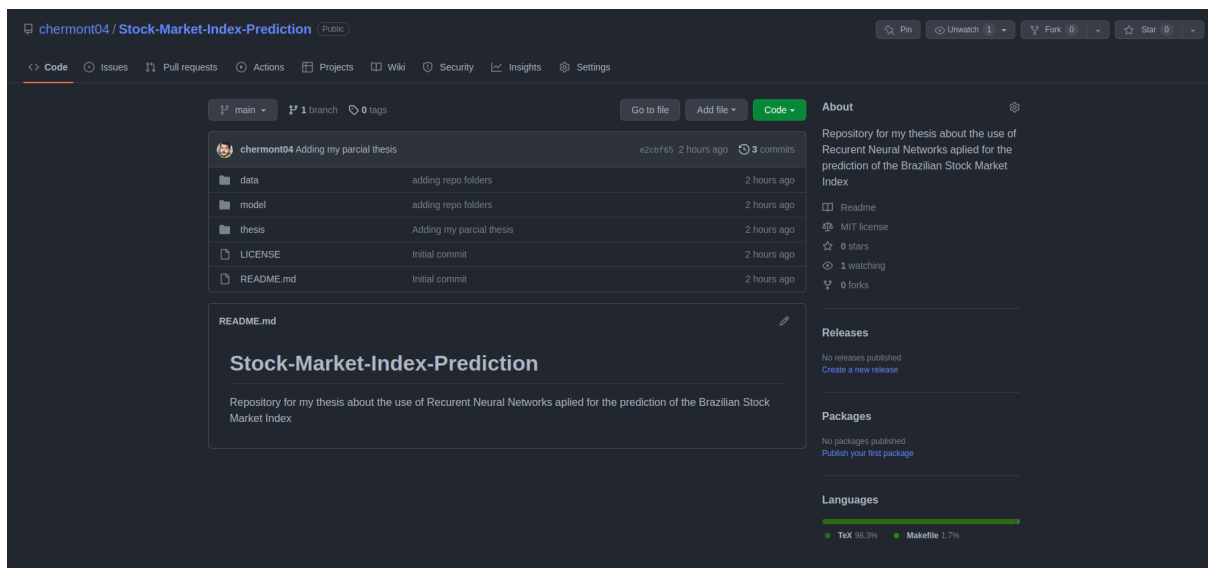
Git é um sistema de controle de versionamento distribuído e gratuito de código aberto. É projetado para lidar com qualquer tipo de projeto e extensões, desde pequenos projetos, até projetos de larga escala. (SCM, 2022)

A ferramenta é capaz de gerir documentos e códigos sem perder seu armazenamento e recuperado em qualquer versão de atualização.

3.1.2 GitHub

O Github é um repositório baseado na web para projetos de software, contando em 2016 com mais de 31 milhões de repositórios. A ferramenta web conta com diversas ferramentas que visam facilitar o desenvolvimento como o rastreamento de problemas, notificações e painéis de status. Utilizando a ferramenta de versionamento de código do Git e a interface gráfica do Github, é possível termos o melhor controle e gestão do projeto. Conforme mostrado na [Figura 8](#) temos a estrutura do projeto, com as pastas referentes ao modelo na *model*, os dados na pasta *data* e a documentação do trabalho de conclusão de curso na pasta *thesis*.

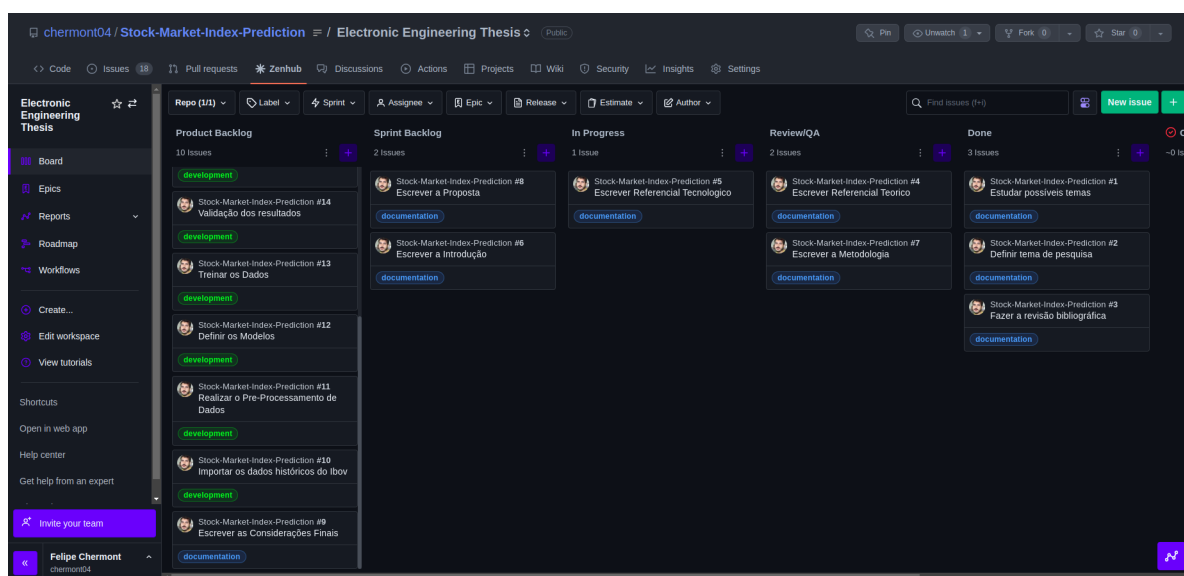
Figura 8 – Repositório da Monografia no GitHub



3.1.3 ZenHub

ZenHub é um *plug-in* conectado diretamente ao GitHub, que torna possível fazer o gerenciamento de tarefas. O quadro facilita visualizar o que precisa ser feito e o progresso realizado. Sendo utilizado para visualizar todas as tarefas dentro do escopo do projeto e seu andamento, como mostrado na Figura 9

Figura 9 – Listagem de Tarefas com o ZenHub



Onde, a primeira coluna é o *Product Backlog* em que estão todas as tarefas necessárias para realização do produto, ou neste caso, da monografia. Na coluna de *Sprint Backlog* estão todas as tarefas que foram realizadas no período de duas semanas, onde são realizadas as entregas parciais do projeto. Na coluna *In Progress* está a tarefa sendo

realizada no momento de visualização do quadro. A *Review/QA* está tudo que já foi finalizado e está em processo de revisão final. Por fim, a coluna de *Done* estão todas as tarefas que foram efetivamente concluídas.

3.2 Desenvolvimento Técnico

Essa seção refere-se às ferramentas e tecnologias utilizadas no desenvolvimento do algoritmo de predição do índice Bovespa, com o uso de redes neurais recorrentes.

3.2.1 Jupyter Notebook

O Jupyter Notebook é uma das ferramentas Jupyter, um software livre, de fácil instalação e com interface gráfica para uma vasta quantidade de linguagens de programação. A Interface do Jupyter Notebook permite criar e compartilhar documentos computacionais, de forma simples e com saídas interativas, tendo uma saída para cada célula, podendo armazenar imagens, textos em HTML, vídeos ou LaTeX em meio ao código, para facilitar a compreensão do usuário e transmitir melhor a mensagem. (JUPYTER, 2022)

3.2.2 Python 3.10.4

Python é uma linguagem de programação com sintaxe clara e concisa, o que facilita a legibilidade do código-fonte, tornando a linguagem mais produtiva. A linguagem conta com diversas estruturas de alto nível (listas, dicionários, tuplas, entre outras) e suporta programação modular, funcional e orientada a objetos (BORGES, 2014). Existem diversos módulos que auxiliam o desenvolvimento nesta linguagem dependendo do objetivo de uso, desde criação de interfaces web a aplicações de inteligência artificial. Dentre a enorme quantidade de pacotes disponíveis, pode-se destacar no uso do desenvolvimento deste trabalho os pacotes [Pandas 1.4.3](#), [Numpy 1.23.0](#), [Matplotlib 3.5.2](#) e o [Keras e TensorFlow 2](#).

3.2.2.1 Pandas 1.4.3

Pandas é uma ferramenta *open source* bastante utilizada em análise de dados, sendo rápida, poderosa, flexível e de fácil utilização, como também é uma das principais bibliotecas Python. A ferramenta é principalmente utilizada para facilitar a leitura e manipulação de dados, tratamento de dados nulos, união de *datasets* entre outras funcionalidades (PANDAS, 2022) que auxiliam toda a etapa de pré-processamento dos dados.

3.2.2.2 Numpy 1.23.0

Numpy é o pacote fundamental do Python para computação científica, sendo rápido e versátil, facilita a vetorização, indexação e transmissão de dados. Oferece diversas funções matemáticas, como geração randômica de números, rotinas de álgebra linear, transformadas de Fourier, entre outras. Por fim o Numpy é um software livre, sendo mantido por uma comunidade no GitHub, sendo possível utilizá-lo em qualquer projeto. (NUMPY, 2022)

3.2.2.3 Matplotlib 3.5.2

O Matplotlib é uma biblioteca utilizada para visualização de dados, sendo possível criar visualizações tanto estáticas quanto animadas em Python e sendo possível exportar em diversos formatos. Facilitando a visualização dos resultados e exibição dos principais resultados do projeto. (MATPLOTLIB, 2022)

3.2.2.4 Keras e TensorFlow 2

Keras é uma *Application Programming Interface* (API) que facilita o uso de *Deep Learning* em projetos Python. A API é construída em cima do TensorFlow 2, uma plataforma de código aberto para *Machine Learning* com diversas ferramentas que permitem o desenvolvimento e implantação de algoritmos de *Machine Learning*, sendo a principal biblioteca para a construção da Rede Neural. O tensorflow possui funções e estruturas de redes neurais prontas, que podem ser invocadas e configuradas da melhor forma possível para o projeto em questão (KERAS, 2022).

Os principais pacotes utilizados neste trabalho desta biblioteca são:

- *Model* para criação do modelo;
- *Sequential* para adicionar as camadas em sequência no modelo inicializado;
- *Dense* para criar a camada de predição;
- *GRU* para adicionar a camada da rede neural GRU;
- *LSTM* para adicionar a camada da rede neural LSTM

3.3 Considerações Finais

Este capítulo apresenta as principais tecnologias e ferramentas utilizadas no desenvolvimento deste trabalho. Descrevendo o uso do Git, GitHub para armazenamento e versionamento de código e o uso do ZenHub para controle e organização de tarefas. Em seguida é explicado como é feita a utilização do Overleaf na escrita científica. Por fim,

é descrito as principais tecnologias utilizadas, sendo o Jupyter notebook para a interação das saídas a cada célula e a mesclagem com as linguagens de escrita e facilidade de compartilhamento. O uso da linguagem Python, pela sua vasta quantidade de pacotes disponíveis e facilidade de compreensão, e os principais pacotes oferecidos que são utilizados no desenvolvimento do trabalho.

4 Metodologia

Este capítulo descreve todas as etapas da metodologia aplicada na elaboração deste trabalho. Primeiramente, há a classificação da [Pesquisa Científica](#), conhecendo a sua natureza de pesquisa, objetivos e procedimentos. Em seguida a [Metodologia de Pesquisa Investigativa](#) que explica as técnicas de pesquisa bibliográfica aplicadas. Na seção de [Metodologia de Desenvolvimento](#) é exemplificado cada atividade e processo seguido na elaboração do trabalho. A [Metodologia de Análise de Resultados](#) descreve de forma sucinta o plano para análise e divulgação de resultados. Já o [Fluxo de Atividades](#) objetivamente demonstra a orientação por etapas para desenvolvimento do trabalho e o fluxo geral de atividades que compreende os Trabalhos de Conclusão de Curso 1 e 2. Por fim, é apresentado o [Cronograma](#) de atividades e as [Considerações Finais](#) do capítulo.

4.1 Pesquisa Científica

Pesquisa pode ser definida como o procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos. (GIL, 1996).

Segundo (GERHARDT; SILVEIRA, 2009) "*A pesquisa científica é o resultado de um inquérito ou exame minucioso realizado com o objetivo de resolver um problema, recorrendo a procedimentos científicos*". Desta forma, a pesquisa pode ser classificada de acordo com diversos tipos. Quanto a **abordagem**, a pesquisa pode ser classificada como qualitativa ou quantitativa. Quanto a **natureza** pode ser classificada como uma pesquisa básica, ou aplicada. Quanto aos **objetivos** temos a pesquisa: exploratória, descritiva ou explicativa. Por fim, quanto aos **procedimentos** a pesquisa pode ser classificada como: experimental, bibliográfica, documental, de campo, ex-post-facto, de levantamento, como *survey*, estudo de caso, participante, pesquisa-ação, etnográfica ou etnometodológica.

4.1.1 Abordagem

A pesquisa quantitativa, tem como foco o pensamento lógico, onde é enfatizado o raciocínio dedutivo, as regras lógicas e os atributos mensuráveis, enquanto a pesquisa qualitativa, procura evidenciar aspectos dinâmicos, holísticos e individuais, para entender o contexto em torno do fenômeno. (GERHARDT; SILVEIRA, 2009)

Desta forma, este trabalho pode ser classificado, quanto a **abordagem** como **quantitativo**.

4.1.2 Natureza

A pesquisa básica objetiva gerar novos conhecimentos, sem uma aplicação prática prevista previamente em seu estudo, envolvendo interesses universais, enquanto a pesquisa aplicada tem como objetivo gerar conhecimentos de aplicações práticas, estudando soluções para problemas específicos, envolvendo interesses específicos. (GERHARDT; SILVEIRA, 2009)

Para o trabalho em questão, pode-se classificar quanto a **natureza** como uma **pesquisa aplicada**.

4.1.3 Objetivos

Segundo (GIL, 1996) os objetivos podem ser classificados como:

- Pesquisa Exploratória: Proporciona maior familiaridade com o problema a partir da construção de hipóteses, envolvendo levantamento bibliográfico, entrevistas com pessoas que tiveram práticas com o problema pesquisado e análise de exemplos.
- Pesquisa descritiva: Descreve fatos e fenômenos de determinada realidade, como estudos de caso, análise documental e pesquisa ex-post-facto.
- Pesquisa Explicativa: Identifica fatores que determinam ou contribuem para a ocorrência dos fenômenos.

Com base nas classificações relatadas, esta pesquisa pode ser classificada como **exploratória**, quanto aos seus **objetivos**.

4.1.4 Procedimento

Conforme citado por (GERHARDT; SILVEIRA, 2009) para desenvolver uma pesquisa, é essencial a escolha do método de pesquisa a se utilizar. Dependendo das características da pesquisa, poderão ser selecionadas diversas modalidades, sendo elas experimental, bibliográfica, documental, de campo, ex-post-facto, de levantamento, como *survey*, estudo de caso, participante, pesquisa-ação, etnográfica ou etnometodológica.

De acordo com as classificações citadas, a **pesquisa bibliográfica** é a mais indicada para este trabalho, quanto ao seu **procedimento**, onde a pesquisa é feita a partir do levantamento de referências teóricas já analisadas e publicadas. É realizado investigações sobre as propostas e análise das diversas posições tomara pelos autores acerca do problema.

Desta forma, a Tabela 2 apresenta o resumo das classificações para o caso de pesquisa deste trabalho, conforme as colocações apresentadas.

Tabela 2 – Resumo de Classificações da Pesquisa Científica

Abordagem	Natureza	Objetivos	Procedimentos
Quantitativa	Aplicada	Exploratória	Pesquisa Bibliográfica

4.2 Metodologia de Pesquisa Investigativa

A pesquisa bibliográfica é desenvolvida a partir de material já elaborado, sendo composto principalmente de livros e artigos científicos. Grande parte dos estudos exploratórios podem ser classificados como pesquisas bibliográficas. (GERHARDT; SILVEIRA, 2009)

4.2.1 *String* de Busca

A definição de uma *string* de busca bem definida é um passo fundamental para o sucesso na condução do estudo, onde impacta de forma significativa a identificação de estudos relevantes e elimina termos que reduzem a precisão da pesquisa ou não agregam estudos relevantes suficientes para o trabalho conduzido. (NAPOLEÃO, 2019)

Seguindo o objetivo principal desta dissertação, desenvolvemos a *string* de busca e seu refinamento, procurando se aproximar da questão de pesquisa: "*É possível a utilização de uma modelagem com base em Redes Neurais Recorrentes na predição do Índice da Bolsa de Valores Brasileira?*" e encontrar de fato os estudos e artigos mais relevantes para elaboração deste estudo. Para a pesquisa foi utilizada a base de dados do Portal de Periódicos da CAPES, que democratiza o acesso a informação científica, reunindo editoras internacionais e nacionais, e atualmente, conta com mais de 49 mil periódicos com texto completo e 455 bases de dados de diversos conteúdos. (CAPES, 2022)

A Tabela 3 retrata a elaboração da *string* de busca, a base de dados e o retorno em quantidade de artigos encontrada para cada *string* utilizada na pesquisa.

Tabela 3 – Elaboração da String de Busca

String	Base de Dados	Quantidade
"Redes Neurais Recorrentes"AND "Bolsa de Valores"	CAPES	14
"Recurrent Neural Network"AND "Stock Market"	CAPES	220
"Recurrent Neural Network"AND "Stock Market" AND "Prediction"	CAPES	164
"Recurrent Neural Network"AND "Stock Market" AND "Prediction"AND "Index"	CAPES	66

4.2.2 Guia em diagramação Prisma

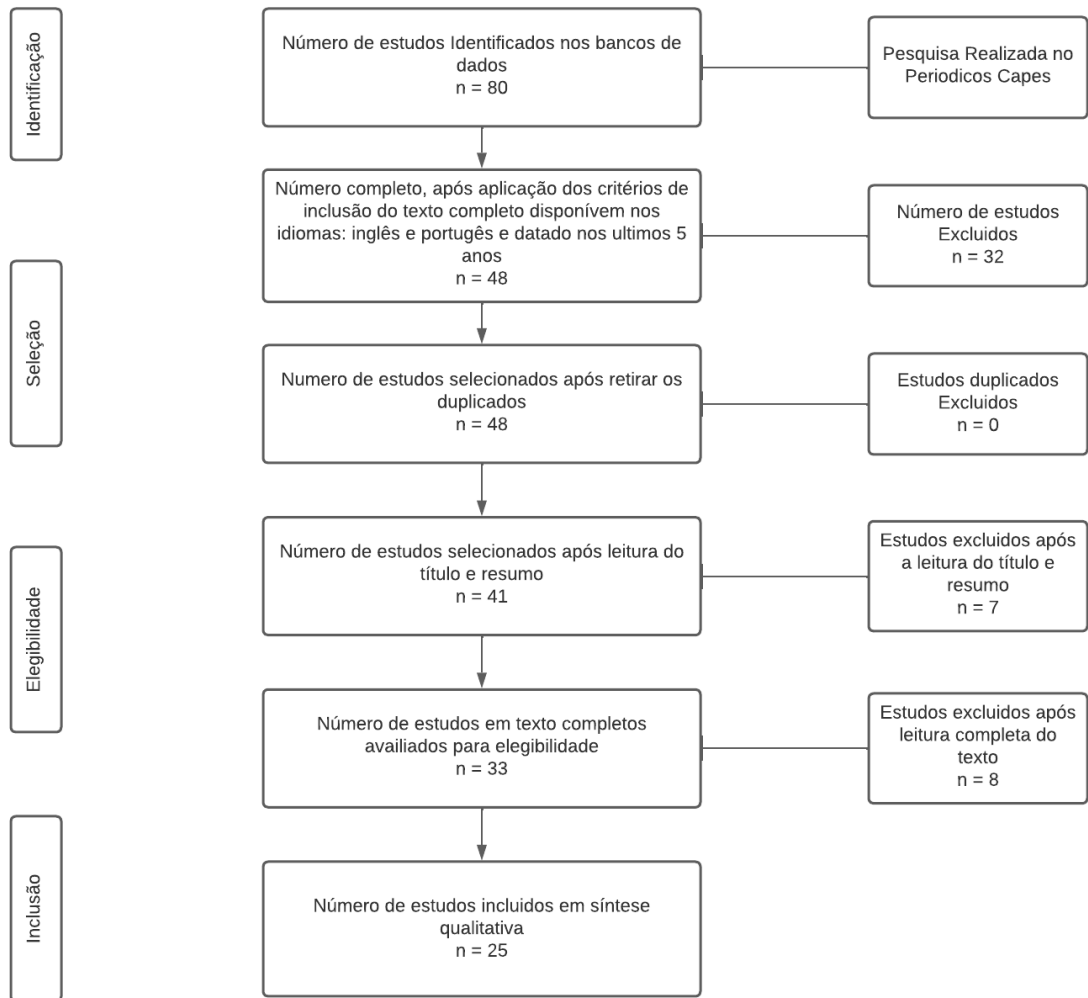
A declaração PRISMA, do inglês (Preferred Reporting Items for Systematic reviews and Meta-Analyses) publicada em 2009 foram projetadas para ajudar autores de revisão sistemáticas a documentar com maior precisão e transparência as etapas e motivações da revisão realizada dos estudos selecionados.(PAGE, 2021)

Desta forma, foi utilizado o Diagrama em Prisma para filtrar os artigos encontrados por nossa *string* de busca de forma a selecionar os artigos mais relevantes para nortear a pesquisa da forma correta e trazer informações mais relevantes possíveis, seguindo as etapas:

- Identificação: Identificar os estudos após o refinamento da *string* de busca.
- Seleção: Filtragem dos artigos duplicados, e filtros que possam ser relevantes dependendo da pesquisa.
- Elegibilidade: Leitura do título e resumo dos artigos selecionados.
- Inclusão: Leitura completa e inclusão dos artigos mais relevantes para seu estudo.

Neste trabalho, após passar por todas as etapas do guia, foram incluídos 25 estudos em síntese qualitativa, conforme mostrado na Figura 10

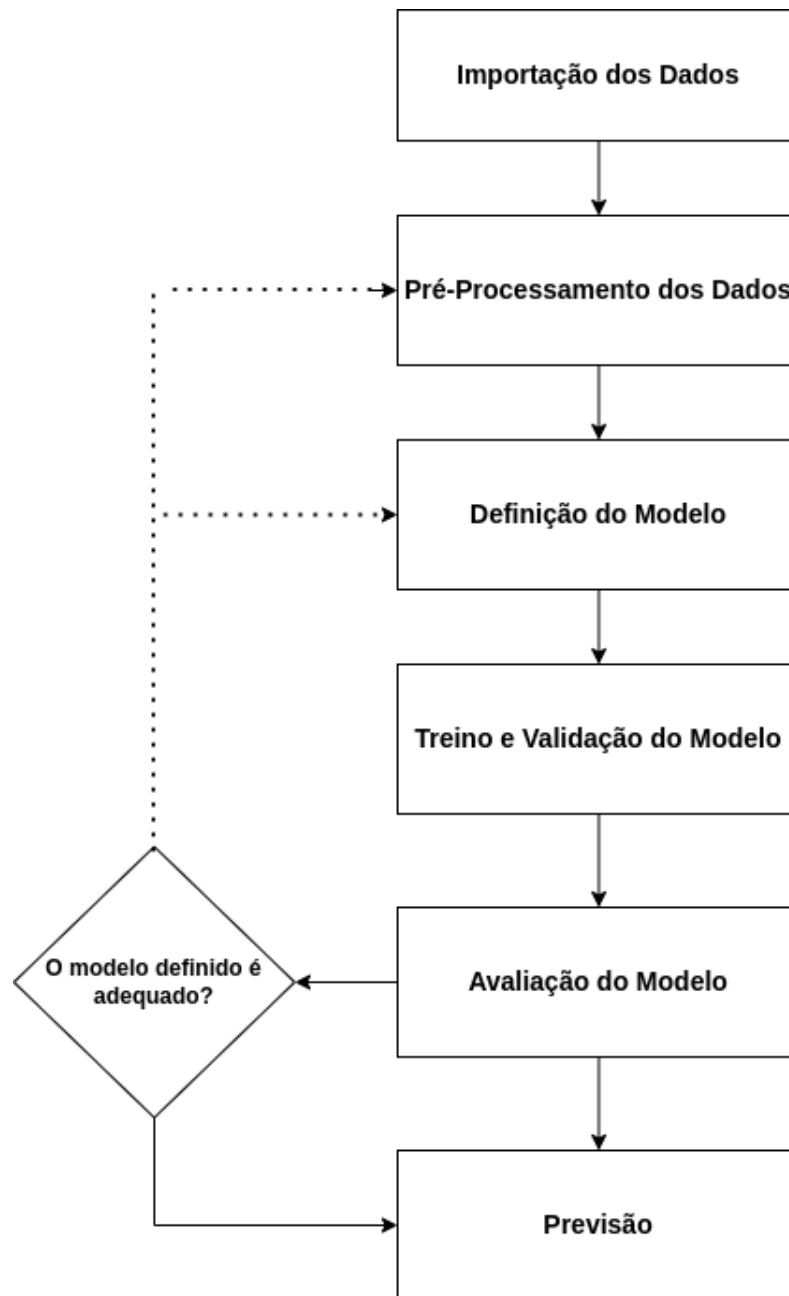
Figura 10 – Diagrama Prisma



4.3 Metodologia de Desenvolvimento

Esta seção detalha as especificações da metodologia de desenvolvimento do algoritmo de predição, sendo classificadas em: importação dos dados, pré-processamento, definição do modelo, treino e validação, avaliação, reajustes necessários e previsão, conforme exemplificado na Figura 11.

Figura 11 – Fluxo de Desenvolvimento do modelo de predição



A importação dos dados, consiste na obtenção dos dados históricos do Ibovespa, para realização das análises da série temporal e desenvolvimento das etapas seguintes.

O pré-processamento é realizado para preparar os dados da forma correta, antes de usá-los como entrada no modelo de inteligência artificial para predição do valor da ação.

A etapa de definição do modelo, será onde iremos definir a arquitetura de rede neural recorrente que será utilizado, definindo os hiperparâmetros explícitos, como o número de épocas e implícitos como o número de camadas, função de ativação e algoritmo de otimização.

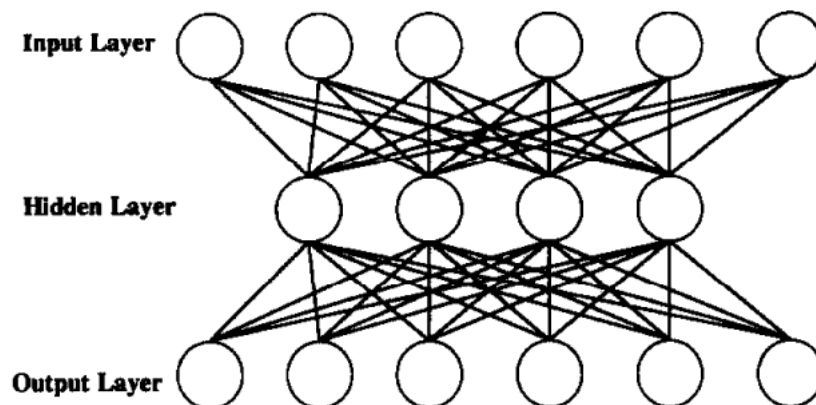
No Treino e Validação do Modelo, é realizado o treinamento do modelo com a técnica do *backpropagation*, que usa uma arquitetura hierárquica em camadas de neurônios simples empregando um alto grau de conectividade entre as camadas.(WYTHOFF, 1993) Desta forma, cada nó recebe uma ou mais entradas da camada anterior e produz uma saída que é transmitida para outro nó de entrada da próxima camada. A Equação 4.1 mostra como é calculado a saída de um determinado nó:

$$\alpha = \sum_{i=1}^n (w_i - x_i) + \theta \quad (4.1)$$

onde x_i é a entrada no nó, w_i representa os pesos aplicados as entradas e θ é o viés de polarização para o nó e n é o número de sinapses para o nó.

É esperado uma rede neural totalmente conectada no final quando se utiliza o método *backpropagation*, conforme a Figura 12.

Figura 12 – Uma rede de *backpropagation* totalmente conectada (WYTHOFF, 1993)



Após a etapa de Treinamento, seguimos para a Avaliação do modelo, onde analisamos as Métricas de Avaliação descritas na Seção 2.3 é realizado todos os ajustes necessários para que seja possível melhorar o modelo até que de fato seja obtido um resultado relevante para a questão de pesquisa.

Por fim, com o modelo treinado e avaliado, é possível realizar a predição do índice, aplicando o modelo para diversos horizontes temporais.

4.4 Metodologia de Análise de Resultados

Como citado por (BUTTON, 2021) além de dominar a metodologia estatística para o planejamento e análise dos dados, o pesquisador deve conhecer bem o que deseja estudar, os dados que irá utilizar e ter uma estimativa qualitativa de como os dados serão analisados.

O físico (MONTGOMERY, 1991) indica um procedimento para planejamento e análise de resultados em etapas, onde:

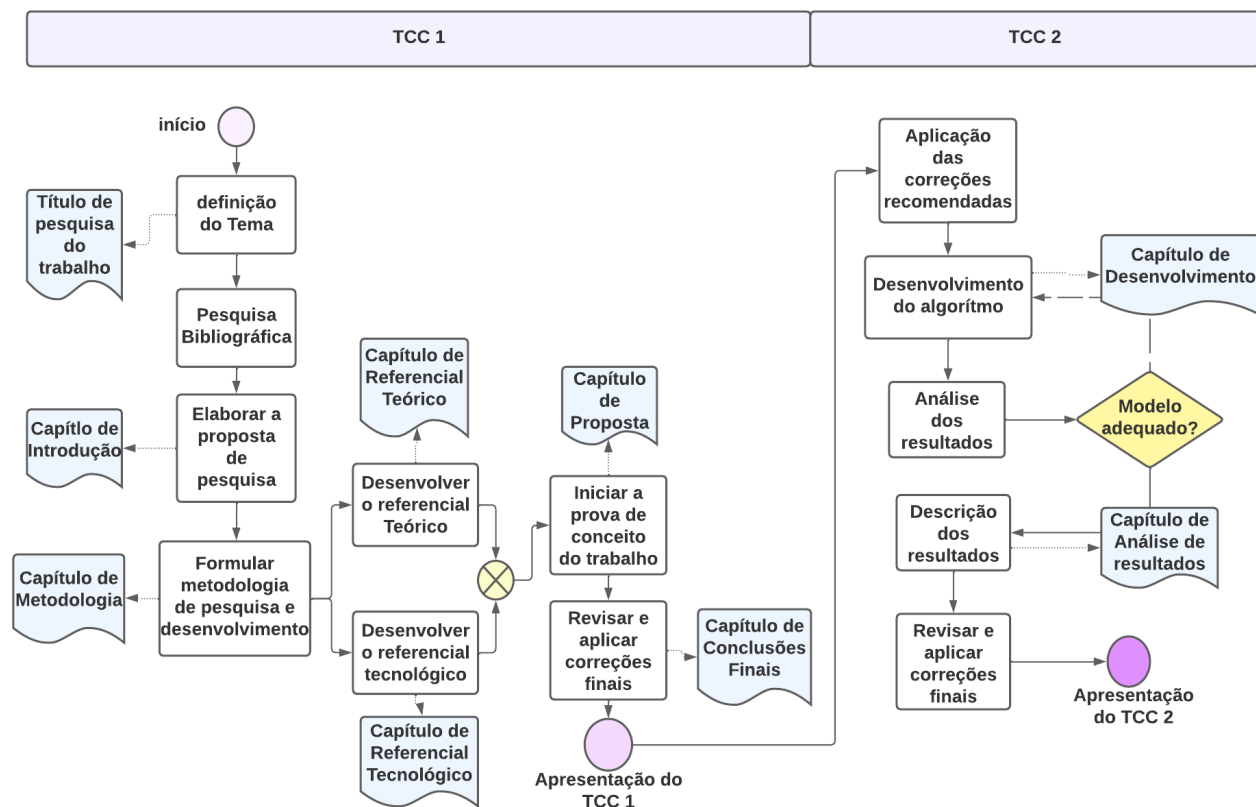
1. Reconhecimento e definição do problema que depende da experiência adquirida anteriormente ou do estudo prévio de processos semelhantes
2. Escolha das variáveis e das faixas de valores que as variáveis serão avaliadas. A avaliação de diversas variáveis pode ser necessária quando o estudo encontra-se em seus estágios iniciais, quando deseja-se verificar uma variável específica, o nível de especificação da avaliação deve ser reduzido, além de manter as demais variáveis influentes em níveis constantes.
3. Escolha adequada da variável de resposta, de modo que seja garantido a objetividade na análise dos resultados. A escolha da variável mais adequada se dá mediante ao erro experimental de medida da variável de resposta seja mínimo, permitindo a análise estatística dos dados.
4. Delineamento dos experimentos: Tamanho da amostra, sequência de execução dos ensaios, necessidade de aleatorização ou uso de blocos. Aprendendo sobre o processo, se deve ser acrescentado novas variáveis ou eliminar variáveis não influentes. Com essa iniciativa, pode-se reduzir o número de ensaios.
5. Execução dos experimentos, monitorando-os e controlando-os. Etapa que garante a validade experimental, onde exige do pesquisador um conhecimento aprofundado dos métodos de controle e monitoramento aplicados.
6. Análise dos resultados, com o uso de métodos estatísticos, para que as conclusões estabelecidas sejam objetivas e verdadeiras, garantindo a confiabilidade e a validade dos resultados, de modo que seja possível determinar o erro associado nas conclusões de acordo com o grau de confiança estabelecido.
7. Elaboração das conclusões e recomendações a partir da análise de resultados. Permitem que decisões sejam tomadas a respeito do processo em estudo, com base em uma documentação, onde pode ser utilizado gráficos e tabelas que apresente de forma clara os resultados obtidos.

4.5 Fluxo de Atividades

O estudo e execução desse trabalho é orientada por etapas.

A Figura 13 apresenta o fluxograma de atividades seguido, onde é apresentado as etapas e atividades essenciais para a condução desse estudo, contemplando os trabalhos de monografia 1 e 2, e suas apresentações finais respectivas.

Figura 13 – Fluxograma de Atividades da Monografia



Definição do Tema: Atividade atribuída a discussão inicial do que o trabalho deveria abordar e refinamento de ideias levantadas através do estudo de redes neurais recorrentes e suas aplicações.

Pesquisa Bibliográfica: Processo de levantamento e coleta de materiais com foco na temática definida, conduzida conforme exemplificado na Seção [Metodologia de Pesquisa Investigativa](#)

Elaboração da proposta de pesquisa: Etapa que descreve o problema em questão, justificativa, objetivos e contextualização do trabalho proposto.

Formulação da metodologia de pesquisa e desenvolvimento: Atividade que definiu o tipo de pesquisa que este trabalho realizará, com base na abordagem, natureza, objetivos e procedimentos, assim como a especificação da metodologia de desenvolvimento do algoritmo, fluxo de atividades e cronograma de elaboração do trabalho.

Desenvolvimento do referencial teórico: Com base nos principais trabalhos e estudos encontrados em nossa pesquisa bibliográfica, exemplificar os principais conceitos abordados no presente trabalho e com base nas outras pesquisas, entender o que será trabalhado e as métricas que devem ser observadas e relatadas.

Desenvolvimento do referencial tecnológico: Esta atividade expõe todas as

principais tecnologias e ferramentas que são utilizadas na elaboração da pesquisa e desenvolvimento da solução proposta.

Início da prova de conceito do trabalho: Esta atividade procura iniciar a proposta de solução, trazendo uma análise inicial do que será abordado futuramente no trabalho de conclusão de curso 2.

Revisão e aplicação das correções finais: Releitura, revisão e aplicação de todas as possíveis correções ao trabalho.

Apresentação do TCC 1: Apresentação a banca julgadora, o trabalho proposto.

Aplicação das correções recomendadas: Após a apresentação, entender as possíveis melhoras e correções apontadas pela banca de forma a termos o melhor trabalho possível.

Desenvolvimento dos algoritmos de predição: Etapa de desenvolvimento da solução proposta inicialmente.

Análise e resultados: Análise dos resultados e comparativo em relação aos principais estudos acerca do tema abordado, conforme descrito da Seção [Metodologia de Análise de Resultados](#).

Descrição dos resultados: Descrição dos principais resultados obtidos e validados à segunda parte da Monografia.

Revisão e aplicação das correções finais: Esta tarefa consiste na revisão de toda a pesquisa e aplicação das últimas correções necessárias, finalizando as conclusões finais do trabalho realizado.

Apresentação do TCC 2: Apresentação final da segunda e última parte da monografia a banca julgadora.

4.6 Cronograma

A Tabela 4 representa o cronograma previsto para as conclusões das tarefas pertencentes ao Trabalho de Conclusão de Curso (TCC) 1, que foi iniciado em Junho de 2022 e tem previsão de conclusão e apresentação ao final do mês de Setembro de 2022.

Tabela 4 – Cronograma TCC 1

Atividade	Jun/22	Jul/22	Ago/22	Set/22
Definição do Tema	X			
Pesquisa Bibliográfica		X		
Introdução			X	
Referencial Teórico			X	
Referencial Tecnológico			X	
Metodologia			X	
Proposta Inicial				X
Conclusões Finais				X
Revisões e finalização da Monografia				X
Apresentação				X

A Tabela 5 trata as atividades previstas na elaboração do Trabalho de Conclusão de Curso (TCC) 2, que iniciará em Outubro de 2022 e está previsto para finalizar em Fevereiro de 2023.

Tabela 5 – Cronograma TCC 2

Atividade	Out/22	Nov/22	Dez/22	Jan/23	Fev/23
Aplicar Correções Sugeridas	X				
Análise, limpeza e filtragem dos dados		X			
Desenvolvimento do Algoritmo de Predição			X		
Análise dos Resultados			X		
Otimizações necessárias no modelo				X	
Descrição dos Resultados				X	
Finalização da escrita da Monografia				X	
Aplicar Correções Necessárias					X
Apresentar					X

4.7 Considerações Finais

Este capítulo apresentou os conceitos de pesquisa científica, classificando este estudo quando a abordagem, natureza, objetivos e procedimentos. Definimos a metodologia de pesquisa investigativa, exemplificando os resultados da construção da *string* de busca e da utilização do guia em diagramação Prisma para obtermos os melhores estudos possíveis correlatos ao tema proposto. Na metodologia de Desenvolvimento, definimos o fluxograma para desenvolvimento do algoritmo de inteligência artificial para predição do Ibovespa, exemplificando as principais técnicas que são utilizadas no treino e validação do modelo. A metodologia de Análise de Resultados indica um procedimento para planejamento de análise de resultados. É apresentado então, o Fluxo de Atividades, onde é listado e exemplificado em ordem de execução e detalhamento cada atividade realizada e a ser realizada para conclusão da monografia. Por fim, tem-se o Cronograma de execução das principais atividades que englobam o Trabalho de conclusão 1 e 2 respectivamente.

5 Proposta

Este capítulo tem como objetivo expor de forma detalhada a proposta deste trabalho conforme apresentado na seção [Metodologia de Desenvolvimento](#). É realizada a [Coleta de Dados](#) que são utilizados na modelagem e o [Pré-Processamento](#) que descreve todos os processos realizados para preparar os dados da melhor forma. A etapa de [Modelagem](#) apresenta a construção dos modelos, e por fim tem-se as [Considerações Finais](#) do capítulo.

5.1 Coleta de Dados

Os dados do preço histórico do IBOVESPA é mapeado pela ação BOVA11 e pode ser coletado através do site do Yahoo! Finanças e baixadas no formato *comma separated value* (csv) ([YAHOO](#),). O conjunto de dados extraídos contém dados de 2 de Janeiro de 2009 a 2 de Fevereiro de 2023, totalizando 3497 registros, possuindo informações de data, valor de abertura, valor máximo, valor mínimo, valor de fechamento, valor de fechamento ajustado e volume transacionado, conforme mostrado na [Figura 14](#).

Figura 14 – Exemplo do conjunto de dados (BOVA11) extraídas do site ([YAHOO](#),), mostrando as 10 primeiras linhas da base de dados

	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
0	2009-01-02 00:00:00-02:00	38.279999	40.320000	37.939999	40.320000	119900	0	0
1	2009-01-05 00:00:00-02:00	40.200001	42.000000	39.700001	42.000000	341700	0	0
2	2009-01-06 00:00:00-02:00	42.060001	42.320000	41.549999	42.320000	178300	0	0
3	2009-01-07 00:00:00-02:00	41.750000	41.770000	40.450001	40.639999	73400	0	0
4	2009-01-08 00:00:00-02:00	40.150002	42.200001	40.150002	42.200001	195300	0	0
5	2009-01-09 00:00:00-02:00	41.980000	42.470001	41.160000	42.000000	271900	0	0
6	2009-01-12 00:00:00-02:00	40.709999	40.750000	39.380001	39.700001	37700	0	0
7	2009-01-13 00:00:00-02:00	0.000000	40.259998	38.540001	40.000000	177300	0	0
8	2009-01-14 00:00:00-02:00	39.400002	39.470001	37.700001	38.029999	205500	0	0
9	2009-01-15 00:00:00-02:00	37.889999	39.360001	36.930000	39.360001	120500	0	0

Os dados presentes na Figura 14 podem ser descritos como:

- **Date (Data)**: data do pregão em que ocorreram as transações;
- **Open (Abertura)**: preço da ação na abertura do mercado, no dia em que iniciaram as negociações;
- **High (Máxima)**: valor mais alto que a ação atingiu durante o pregão do dia;
- **Low (Mínima)**: menor valor que a ação atingiu durante o pregão do dia;
- **Close (Fechamento)**: valor da ação no momento do fim do pregão;
- **Volume (Volume)**: quantidade de ações negociadas durante o pregão.

Entre as variáveis de preço informadas acima, o preço de fechamento é sempre o mais preocupante pelos investidores (ZHENG; WANG; CHEN, 2020). Portanto, será utilizado o preço de fechamento do pregão como variável de predição.

5.2 Pré-Processamento

O pré-processamento é uma etapa importante, no qual transforma-se o dado em um formato compreensível, de forma que seja facilmente analisado e interpretado pelo algoritmo. Dados não processados geralmente acabam levando os algoritmos a terem uma modelagem ruim e baixa acurácia (SAKO BERTHINE NYUNGA MPINDA, 2022).

5.2.1 Análise Exploratória

Inicialmente são retiradas as linhas onde o volume é zero, pois não trazem nenhuma informação relevante para o treinamento. Em seguida os dados são ordenados pela data, da mais antiga para a mais recente, a fim de organizar os dados para o aprendizado. Podemos verificar a variação do valor da ação no fechamento do pregão ao longo do tempo, conforme a Figura 15.

Figura 15 – Preço no Fechamento do Pregão do BOVA11 de 2009 a Fevereiro de 2023



5.2.2 Engenharia de Variáveis

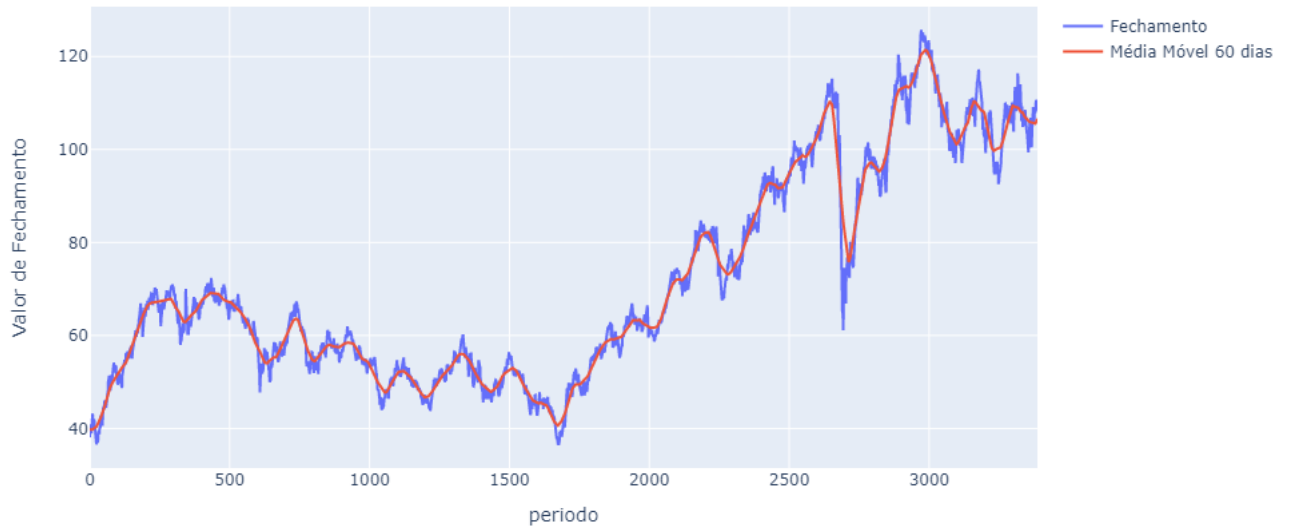
Para melhorar o treinamento das redes neurais, foram adicionadas novas variáveis (*features*) utilizadas nas análises técnicas que auxiliam na predição como as médias móveis, as bandas de Bollinger, o desvio padrão e o índice de força relativa.

A técnica da média móvel é um método adequado quando procura-se ajustar as irregularidades de curto prazo na série de dados (BORELLA et al., 2019) e pode ser calculada de acordo com a equação 5.1

$$MédiaMóvel = \frac{1}{n} \sum_{i=1}^n p_i \quad (5.1)$$

Em que n é referente ao número de períodos considerados e p_i o preço no respectivo período. O ajuste e suavização propostos pela média móvel pode ser visto na 16.

Figura 16 – Demonstrativo do efeito da média móvel na série temporal



As bandas de Bollinger foram desenvolvidas para compreender a volatilidade de mercados capitais, analisando duas médias móveis com desvios padrões superiores e inferiores (NETO, 2018) conforme mostrado nas equações 5.2 e 5.3.

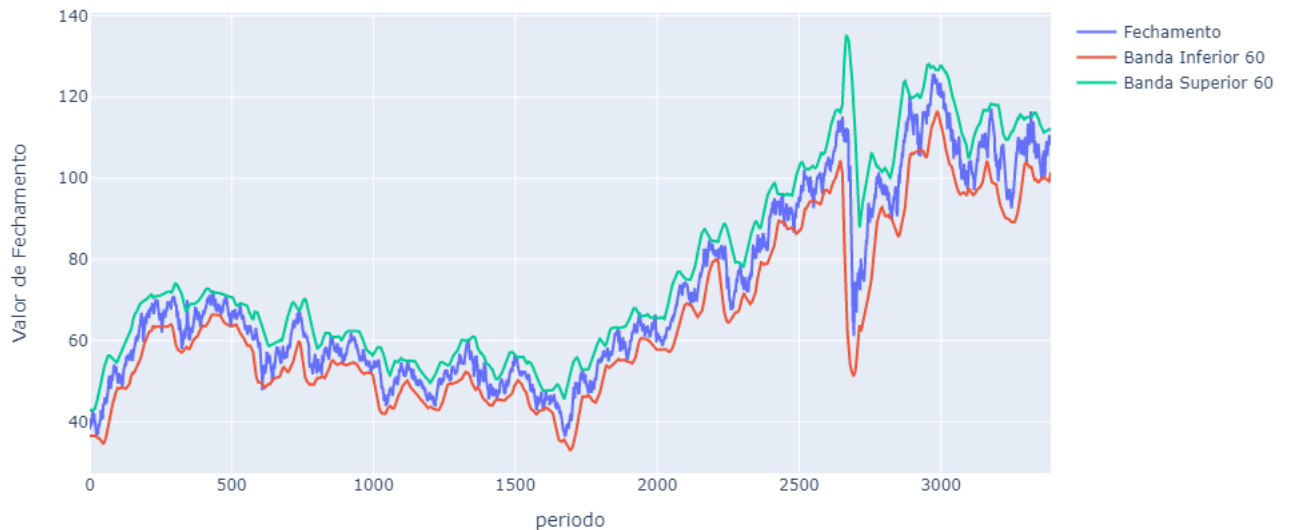
$$BS = MM + (2 * DP) \quad (5.2)$$

$$BI = MM - (2 * DP) \quad (5.3)$$

Onde BS e BI são as bandas de bollinger superiores e inferiores respectivamente, MM se refere as médias móveis e DP representa o desvio padrão.

A Figura 17 mostra como é aplicado os limites das bandas de bollinger dos últimos 60 dias dos valores de fechamento da ação.

Figura 17 – Demonstrativo das bandas de bollinger aplicadas ao longo do tempo da ação do BOVA11



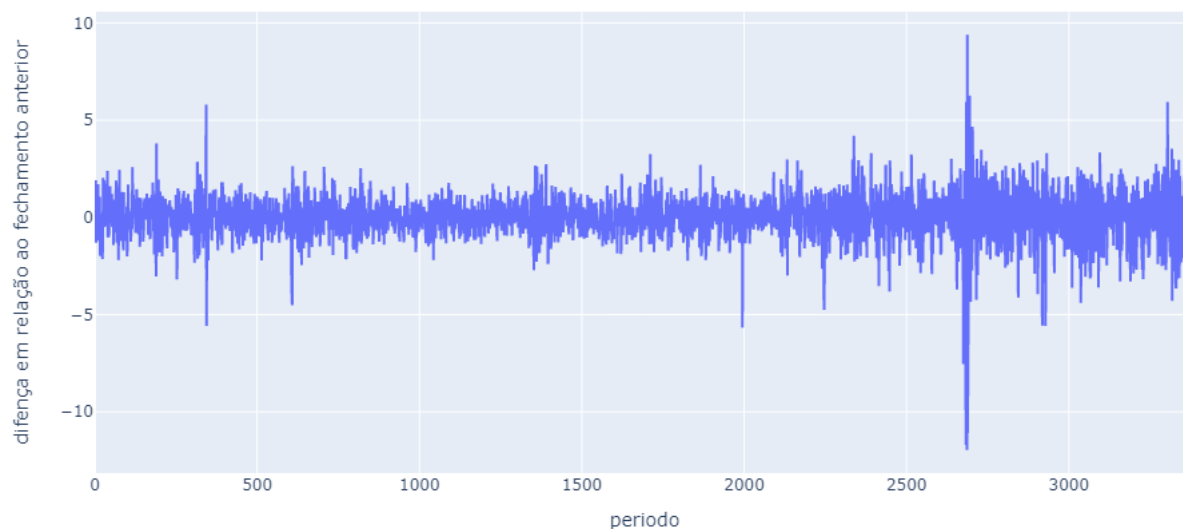
Outro indicador relevante adicionado é o índice de força relativa (IFR), que se comporta como um oscilador, medindo a velocidade e a mudança dos movimentos de preço, e pode ser medido como mostrando na equação 5.4.

$$IFR = 100 - (1 + FR) \quad (5.4)$$

Em que FR é a força relativa, representada pela divisão entre o ganho médio e a perda média.

A última variável adicionada é a diferença entre o valor de fechamento do mercado no dia anterior e o fechamento atual. Desta forma pode-se verificar a volatilidade do índice ao longo dos anos em relação ao dia anterior, como demonstrado na Figura 18.

Figura 18 – Variação do fechamento de mercado em relação ao dia anterior



Conforme apresentado acima, foi adicionado aos dados de entrada dos modelos diversas *features* auxiliares que melhoram a qualidade dos modelos nas previsões futuras, conforme a lista abaixo:

- `close` é o valor de fechamento da ação no dia;
- `close_change` é a diferença entre o valor de fechamento do mercado do dia anterior e o fechamento atual;
- `ma_3d` é a média móvel dos últimos 3 valores de fechamento da ação;
- `ma_10d` é a média móvel dos últimos 10 valores de fechamento da ação;
- `ma_30d` é a média móvel dos últimos 30 valores de fechamento da ação;
- `ma_60d` é a média móvel dos últimos 60 valores de fechamento da ação;
- `std_07d` é o desvio padrão dos últimos 7 valores de fechamento da ação;
- `std_30d` é o desvio padrão dos últimos 30 valores de fechamento da ação;
- `std_60d` é o desvio padrão dos últimos 60 valores de fechamento da ação;
- `UPPER3` é o limite superior da banda de bollinger nos últimos 3 dias;
- `UPPER30` é o limite superior da banda de bollinger nos últimos 30 dias;

- UPPER60 é o limite superior da banda de bollinger nos ultimos 60 dias;
- LOWER3 é o limite inferior da banda de bollinger nos ultimos 3 dias;
- LOWER30 é o limite inferior da banda de bollinger nos ultimos 30 dias;
- LOWER60 é o limite inferior da banda de bollinger nos ultimos 60 dias;
- RSI é o índice da força relativa em um período de 15 dias;

Por fim, os dados são normalizados para valores no intervalo $[-1,1]$ através da função *MinMaxScaler* conforme a Equação 5.5 e separados entre treino e teste, onde 90% dos dados serão utilizados para treino do modelo e 10% será utilizado para testagem. A normalização dos dados faz com que os valores apresentados na entrada da rede não se afastem dos limites das funções de ativação tangente e sigmoide, aproxima a escala dos dados de entrada com os pesos iniciais e aumenta a acurácia dos modelos de aprendizado de máquina. (PASQUOTTO, 2010)

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5.5)$$

5.3 Modelagem

Com os dados tratados e pré-processados, pode-se construir os modelos de rede neural LSTM e GRU propostos. A construção é realizada em linguagem Python utilizando a interface de implementação para algoritmos de aprendizado de máquinas TensorFlow, conforme descritos no Capítulo Referencial Tecnológico.

As modelagens de redes neurais recursivas objetiva aprender com as informações mais relevantes dos indicadores de entrada anteriores e melhorar a previsão. Ambas modelagens são compostas de uma entrada sequencial, seguidas por camadas de LSTM no caso do modelo em *Long Short Term Memory* ou camadas de GRU no caso do modelo em *Gated Recurrent Units*. Entre as camadas é adicionado uma regularização conhecida por *dropout* que derruba uma certa quantidade de neurônios de forma aleatória afim de evitar *overfitting* no modelo de previsão (PAZ, 2019). Para a construção das redes neurais foram importados alguns módulos do Keras:

- *Sequential* para inicializar a rede neural;
- *Dense* para adicionar uma camada de rede neural densamente conectada
- LSTM para adicionar a camada de memória
- GRU para adicionar a camade de *gate* de memória

- *Dropout* para adicionar camadas de ajuste
- *Early stopping* é parada antecipada do treinamento para evitar problemas de *overfitting*

5.3.1 Hiperparâmetros

"Hiperparâmetros são parâmetros cujos valores são utilizados para controlar o processo de aprendizado de redes neurais" (PASQUOTTO, 2010).

A otimização de ambos modelos utilizados foi o otimizador Adam. Segundo (P. BA, 2019) o método combina as vantagens de dois métodos de otimização bem conhecidos, sendo eles o AdaGrad que lida com gradientes esparsos e o RMSProp que lida com objetivos não estacionários. O otimizador Adam também possui uma implementação mais simples, não necessitando de muita memória, o que é ideal para quando há uma grande base de dados ou muitos parâmetros.

Outros parâmetros como o número de camadas, *batch size*, *dropout* e janela de observação passaram por um processo de tunagem, onde foram testados 24 combinações diferentes rodando o modelo com cada combinação e analisando qual combinação apresentava o menor erro quadrático médio (RMSE) em ambos os modelos em que:

- O número de neurônios da rede poderia ter 64, 128 ou 256 unidades;
- *Batch size* poderia ser 32 ou 64 lotes;
- *Dropout* de 0.2 ou 0.1;
- Janelas de tempo de 50 ou 100 dias passados;

De acordo com as tabelas 7 e 6 é identificado que na rede LSTM, será utilizada 1 camada com 256 unidades, 0.1 de *Dropout*, *Batch size* de 32 observando uma janela de 50 dias passados, já para a rede GRU será utilizado 1 camada com 256 unidades, 0.1 de *Dropout*, *Batch size* de 32, observando uma janela de 100 dias passados.

Tabela 6 – Tunagem dos hiperparâmetros da rede GRU

Unidade	Batch Size	Dropout	Janela de tempo	RMSE
256	32	0.1	100	0.0235
256	32	0.2	50	0.0241
256	64	0.1	50	0.0247
256	64	0.1	100	0.0247
128	32	0.2	100	0.0252
128	64	0.2	50	0.0263
128	32	0.2	50	0.0267
128	64	0.1	50	0.0269
256	64	0.2	100	0.0273
256	32	0.1	50	0.0277
128	32	0.1	50	0.0278
64	32	0.1	100	0.0279
64	32	0.2	100	0.0282
64	32	0.2	50	0.0285
64	64	0.1	100	0.0291
128	64	0.1	100	0.0292
64	64	0.2	50	0.0307
256	32	0.2	100	0.0308
64	32	0.1	50	0.0325
64	64	0.2	100	0.0329
64	64	0.1	50	0.0336
128	32	0.1	100	0.0346
128	64	0.2	100	0.0356
256	64	0.2	50	0.0373

Tabela 7 – Tunagem dos hyperparâmetros da rede LSTM

Unidade	Batch Size	Dropout	Janela de Tempo	RMSE
256	32	0.1	50	0,0224
256	32	0.2	100	0,0264
64	64	0.1	100	0,0276
128	32	0.1	100	0,0280
256	64	0.2	100	0,0282
64	32	0.2	50	0,0284
256	64	0.1	50	0,0288
128	32	0.1	50	0,0291
128	64	0.2	100	0,0292
128	64	0.1	100	0,0295
64	32	0.1	50	0,0299
64	32	0.1	100	0,0302
64	64	0.1	50	0,0304
256	64	0.1	100	0,0304
256	64	0.2	50	0,0305
256	32	0.1	100	0,0308
64	32	0.2	100	0,0315
128	64	0.1	50	0,0324
64	64	0.2	100	0,0330
64	64	0.2	50	0,0336
128	64	0.2	50	0,0366
128	32	0.2	50	0,0370
256	32	0.2	50	0,0370
128	32	0.2	100	0,0381

5.3.2 Hardware

Quando se trata de redes neurais, as especificações técnicas de *hardware* podem ser bastante importantes, pois dependendo da complexidade e quantidade de dados, pode-se ter dificuldades na execução desses experimentos. As especificações técnicas da máquina utilizada para rodar todos os processos descritos neste trabalho são descritas na tabela 8:

Tabela 8 – Especificações de Hardware

Hardware	Modelo
Processados	i7 - 5500U 2.4 GHz
RAM instalada	8.00 GB
HD	1 TB

5.4 Considerações Finais

Esse capítulo compreendeu a proposta desse Trabalho de Conclusão de Curso, visando apresentar como é realizada a coleta de dados e a descrição de cada coluna, o pré-processamento dos dados onde realiza-se o tratamento dos dados, adição de informações importantes para o treinamento e normalização dos dados e por fim, na modelagem é apresentado como é realizada a construção do modelo proposto e a definição dos hiperparâmetros.

6 Resultados e Discussões

Foram coletadas 3 métricas de avaliação dos modelos de LSTM e GRU, sendo elas o erro médio absoluto (MAE), o percentual absoluto médio (MAPE) e o erro da raiz média quadrada (RMSE), descritos na seção [Métricas de Avaliação](#). As figuras 19 e 20 abaixo representam os resultados gráficos dos *datasets* de teste em relação ao valor original. Percebe-se que apesar de aparentemente os valores serem bem próximos, em ambos os gráficos, a predição se apresenta um pouco distoada do valor real, mostrando uma variação um pouco menor, e bastante parecido com os valores reais do dia anterior.

Figura 19 – Resultados obtidos no teste da predição com a rede LSTM nas ações do Ibovespa

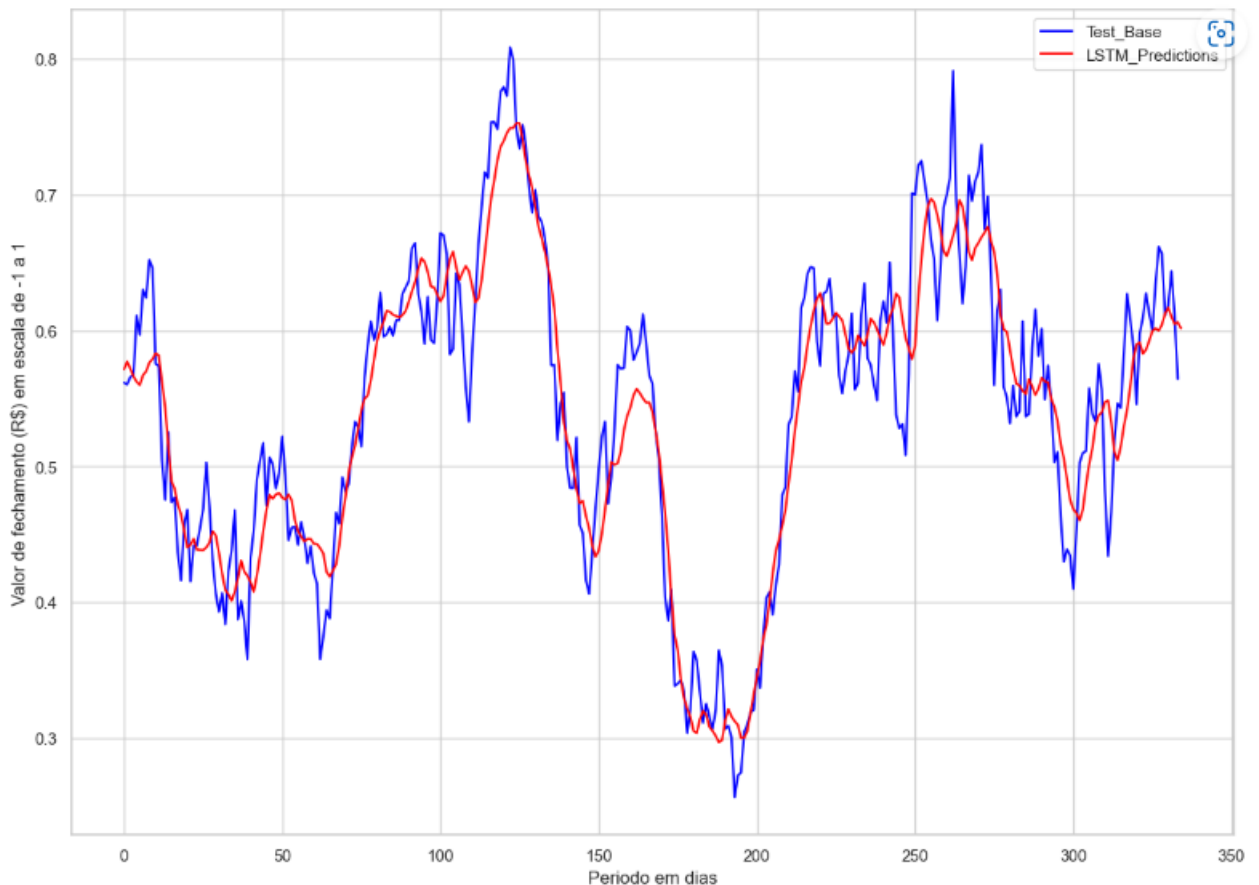
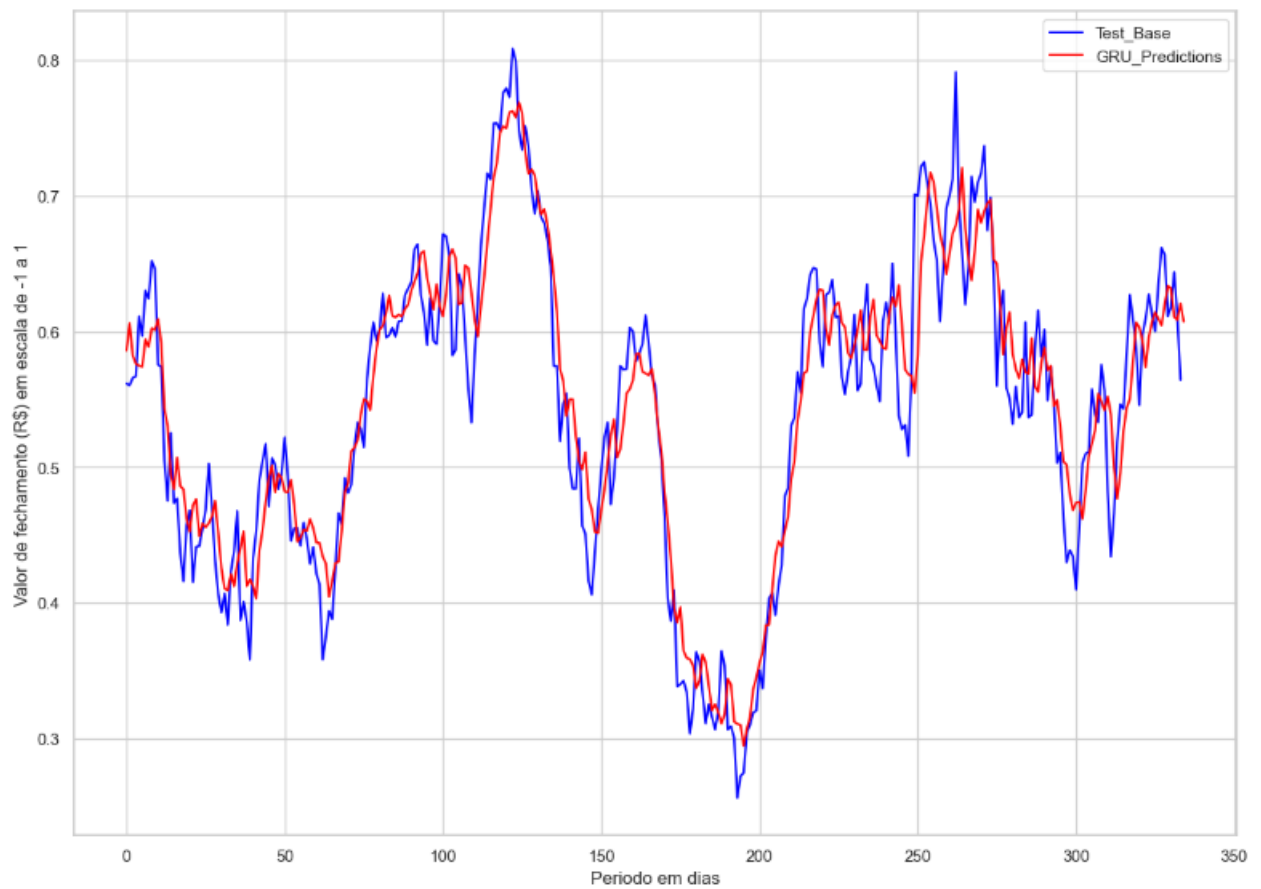


Figura 20 – Resultados obtidos no teste da predição com a rede GRU nas ações do Ibovespa



A tabela 9 apresenta os resultados da predição realizada no *dataset* de teste, onde pode-se perceber que a rede LSTM e GRU se comportaram de forma bem parecida com leve vantagem para a modelagem GRU nas principais métricas de erro avaliadas

Tabela 9 – Comparativo das métricas de erro coletadas dos testes

Rede	MAE	MAPE	RMSE
LSTM	0.032	6.11	0.040
GRU	0.031	6.10	0.038

Comparando os resultados obtidos em relação a pesquisa bibliográfica apresentada na seção [Trabalhos Correlatos](#), onde também são aplicadas técnicas de redes neurais recursivas na predição de um índice mostra que os autores (LV et al., 2022) e (WANG et al., 2022) obtiveram resultados similares com o apresentado neste estudo, como mostrado na tabela 10 abaixo.

Tabela 10 – Comparativo dos resultados analisados na pesquisa bibliográfica e do desenvolvido pelo Autor

Predição	Modelo	Métrica	Resultado	Literatura		
Ibovespa	LSTM	MAPE	6.18	Autor		
		MAE	32.7			
		RMSE	40.8			
	GRU	MAPE	6.11			
		MAE	31.56			
		RMSE	38.68			
Standard and Poor's 500 (S&P500)	LSTM	MAE	33,5	(LV et al., 2022)		
		RMSE	53,4			
	GRU	MAE	43,3			
		RMSE	63,2			
	ARIMA	MAE	34,1			
		RMSE	54,8			
	Standard and Poor's 500 (S&P500)	LSTM	MAE		15,66	(WANG et al., 2022)
			RMSE		23,89	
		RC Model	MAE		15,80	
RMSE			23,25			
RNN		MAE	15,37			
		RMSE	23,44			

Apesar de se tratar de ações e tempos diferentes, o resultado comparativo mostra que as técnicas aplicadas estão em linha com outros estudos seguindo a mesma ótica.

Outro objetivo importante deste estudo é avaliar o comportamento das redes neurais recursivas na predição de valores futuros da ação. Para este objetivo foi segmentado os 90 últimos dias de registro dos dados e observado a predição das redes para uma janela futura de 1, 5, 30 e 90 dias, conforme mostrado na tabela 11.

Tabela 11 – Comparativo das métricas de erro entre os modelos para 1, 5, 30 e 90 dias previsionados

Rede	Dias de predição	MAE	MAPE	RMSE
LSTM	1	0.048	7.86	0.048
GRU	1	0.069	10.95	0.069
LSTM	5	0.026	4.4	0.030
GRU	5	0.019	3.21	0.024
LSTM	30	0.047	8.92	0.053
GRU	30	0.041	7.50	0.051
LSTM	90	0.13	44.74	0.18
GRU	90	0.056	10.17	0.068

Observando as redes LSTM e GRU, percebe-se que a modelagem em GRU apresenta um comportamento melhor ao longo do tempo, enquanto a LSTM apresentou um resultado relativamente melhor apenas na predição em uma janela de 1 dia. O erro é claramente propagado quanto maior a janela de tempo observada, maior a distância entre a predição e o valor real, como pode ser observado nas Figuras 21, 22, 23 e 24 para 1, 5, 30 e 90 dias respectivamente.

Figura 21 – Comparativo entre as arquiteturas para 1 dia previsionado

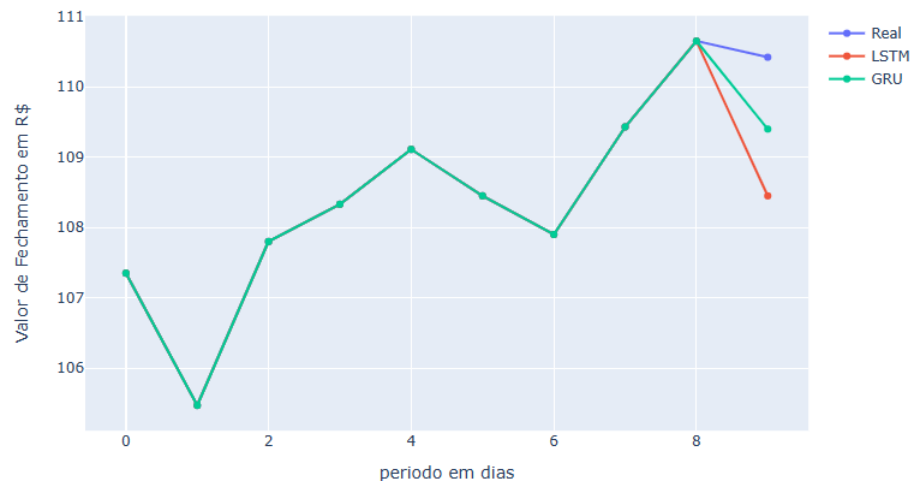


Figura 22 – Comparativo entre as arquiteturas para 5 dia previsionado

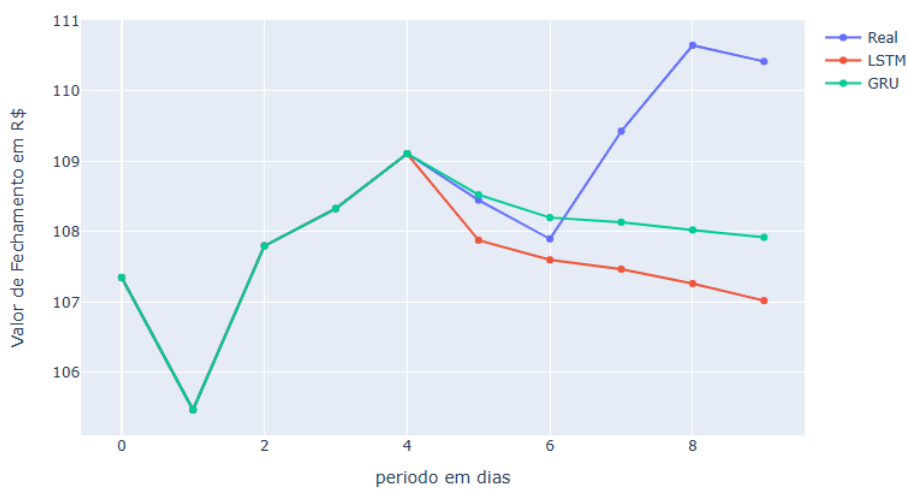
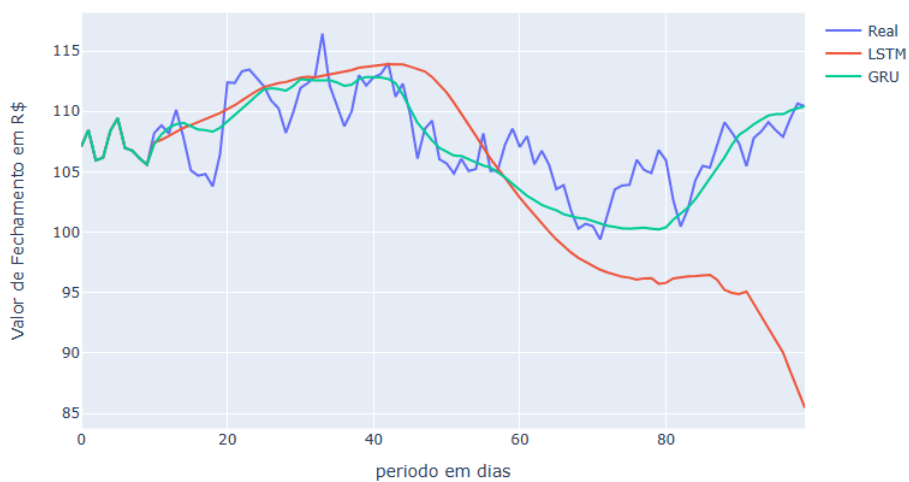


Figura 23 – Comparativo entre as arquiteturas para 30 dia previsionado



Figura 24 – Comparativo entre as arquiteturas para 90 dia previsionado



Pelo comportamento gráfico comparativo entre as arquiteturas e o valor real de fechamento da ação, nota-se que as redes neurais tem um comportamento bem mais suavizado, não capturando realmente as grandes variações que o mercado acaba enfrentando, porém a rede GRU conseguiu se adaptar relativamente bem as nuâncias do mercado até períodos mais longos tendo resultados satisfatórios quando comparado com a rede LSTM. Nas Figuras 25, 26, 27 é representado a diferença entre o valor do fechamento da ação em determinado dia e o valor de fechamento anterior para 5, 30 e 90 dias respectivamente.

Figura 25 – Comparativo da variação entre dias para 5 dias previsionados

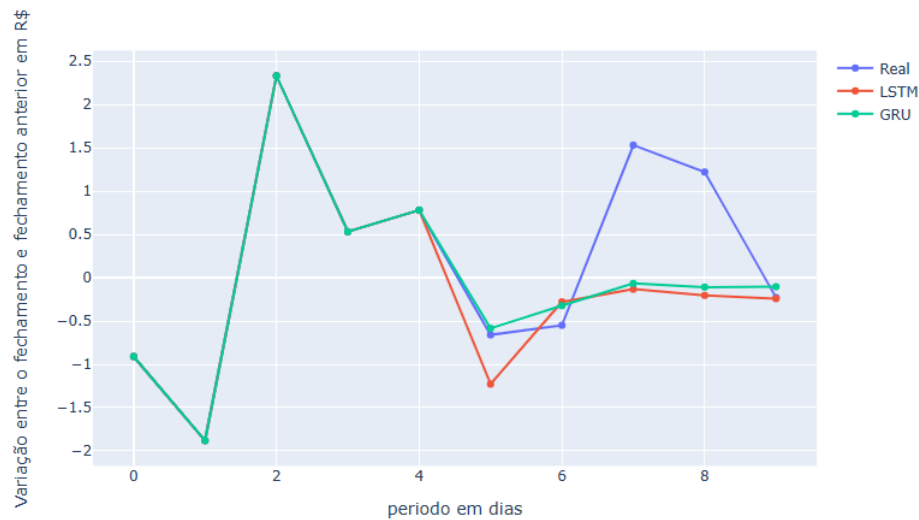


Figura 26 – Comparativo da variação entre dias para 30 dias previsionados

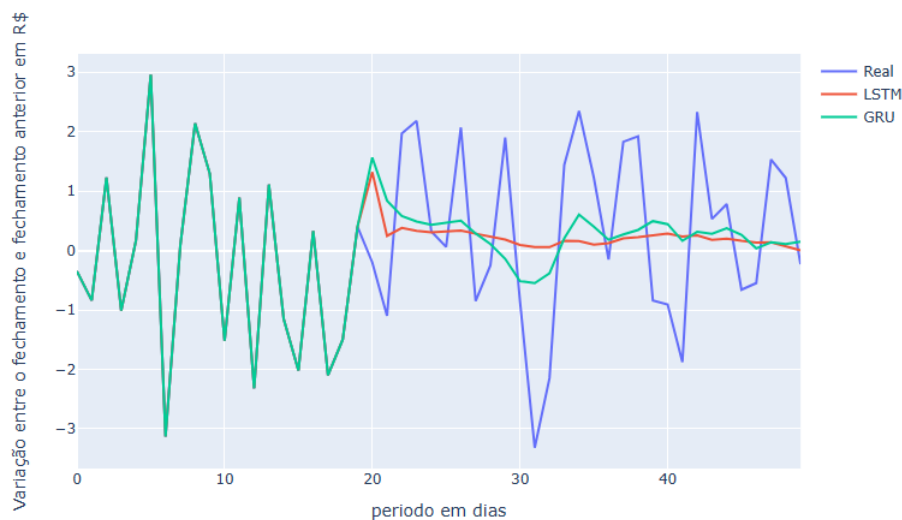
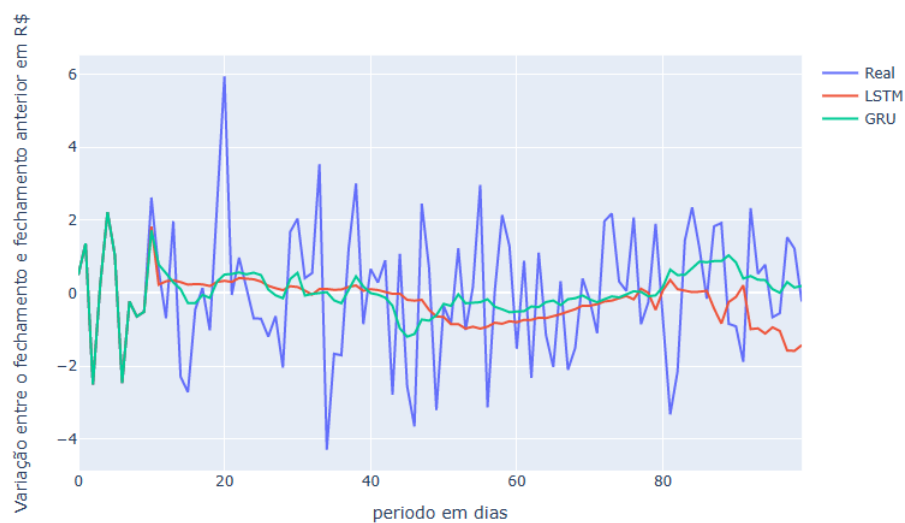


Figura 27 – Comparativo da variação entre dias para 90 dias previsionados



Pela representação gráfica, é possível perceber claramente que enquanto o mercado oscila bastante entre os valores de fechamento ao longo dos dias, as redes neurais tem um comportamento bastante linear, não sendo capazes de mapear a mesma oscilação de forma precisa.

7 Considerações Finais

Este capítulo objetiva apresentar as conclusões e os resultados obtidos com base na proposta deste trabalho. Na seção [Status do Trabalho](#) apresentamos um panorama das principais atividades realizadas referentes ao Trabalho de Conclusão de Curso e o *status* referente aos objetivos apresentados na seção de [Objetivos](#).

A predição da bolsa de valores é um tema amplamente abordado e novas técnicas e metodologias são aplicadas com frequência a fim de tentar chegar o mais próximo possível do valor de uma ação. A avaliação de técnicas de redes neurais recursivas tem potencial por se tratar de um modelo de rede bastante utilizado em séries temporais e por sua alta capacidade de aprendizado com valores passados, podendo auxiliar o investidor a tomar decisões mais corretas com o auxílio de inteligência artificial.

No intuito de contribuir com esse cenário, buscou-se com esse trabalho responder a seguinte Questão de Pesquisa:

É possível utilizar redes neurais recorrentes para prever o índice da bolsa de valores brasileira?

Diante dos resultados apresentados, em particular no Capítulo [Resultados](#), pode-se concluir que as técnicas apresentadas neste trabalho de LSTM e GRU ainda não são capazes de realizar uma predição acurada do valor de fechamento do índice da bolsa de valores, porém podem indicar, no curto prazo uma tendência mais conservadora.

Percebeu-se que, apesar dos resultados não serem muito satisfatórios na predição do índice, a rede LSTM obteve resultados mais próximos, sendo possivelmente uma modelagem melhor para seguimento em estudos e arquiteturas mais avançadas acerca deste tema.

7.1 Status do Trabalho

Com base nas seções de [Fluxograma de Atividades da Monografia](#) e [Cronograma](#), pode-se aferir as atividades realizadas neste Trabalho de Conclusão de Curso. A tabela [12](#) informa o status dessas atividades quanto a conclusão ou não realização de cada uma para o Trabalho de Conclusão de Curso 1, enquanto a tabela [13](#) informa os status referentes ao Trabalho de Conclusão de Curso 2.

Tabela 12 – *Status* das atividades referentes ao TCC 1

Atividade	Status
Definição do Tema	Concluída
Introdução	Concluída
Referencial Teórico	Concluída
Metodologia	Concluída
Proposta Inicial	Concluída
Finalização da Monografia	Concluída
Revisões Finais	Concluída
Apresentar	Concluída

Tabela 13 – *Status* das atividades referentes ao TCC 2

Atividade	Status
Aplicação das correções sugeridas	Concluída
Estudo e refinamento de variáveis	Concluída
Tunagem dos hiperparâmetros	Concluída
testagem do modelo	Concluída
Realização das previsões	Concluída
Finalização da Proposta	Concluída
Finalização dos Resultados	Concluída
Apresentar	Concluída

Em relação aos objetivos específicos apresentados na seção de [Objetivos](#) pode-se aferir:

- Avaliar os dados históricos de volume negociado, valor na abertura, valor de fechamento, máxima no dia e mínima no dia, do IBOVESPA através de ferramentas estatísticas:

Status: Importação e Avaliação dos dados realizadas, conforme apresentado na seção [Coleta de Dados](#).

- Verificar o efeito qualitativo das técnicas de pré-processamento:

Status: Realizado o estudo de variáveis e técnicas que poderiam contribuir positivamente para a previsão, conforme adicionadas ao modelo conforme a seção [Pré-Processamento](#) podendo ainda ter diversos ajustes conforme avanço do projeto e novos testes com diferentes modelos.

- Realizar um estudo dos hiperparâmetros, suas influências na rede e a melhor combinação possível para cada modelo;

Status: Realizado o processo de tunagem dos hiperparâmetros conforme descrito na seção de [Modelagem](#) variando as unidades, *batch size*, *dropout* e a janela de

tempo. Podendo passar por ajustes aumentando a quantidade de hiperparâmetros e combinações para testagem.

- Avaliar a qualidade das previsões num espaço de tempo de 5, 30, 60 e 90 dias a frente:

Status: Realizado as previsões e comparativo entre os modelos LSTM e GRU conforme descrito no capítulo [Resultados](#).

- Comparar as técnicas e resultados obtidos com a literatura analisada na pesquisa bibliográfica:

Status: Realizado no comparativo dos resultados obtidos pelos testes em ambos modelos obtidos neste trabalho no capítulo [Resultados](#) e nos trabalhos apresentados na seção [Trabalhos Correlatos](#)

Os objetivos específicos propostos nesse trabalho foram alcançados conforme descritos na seção [Objetivos Específicos](#). Foram levantados as principais aplicações de redes neurais recorrentes aplicadas no mercado de ações. Com base neste referencial teórico, foi possível elaborar arquiteturas em LSTM e GRU e finalmente, avaliar o comportamento destes modelos na tentativa da predição do índice da bolsa de valores brasileira.

7.1.1 Trabalhos Futuros

O trabalho apresentado contém diversos pontos de melhoria, como o refinamento das técnicas de pré-processamento, que avalia possíveis indicadores técnicos que possam contribuir na predição, avaliação de *outliers* e entre outros. Na etapa de modelagem existem diversas combinações de hiperparâmetros que podem ser testadas, a fim de se construir o modelo de forma mais próxima possível.

Além do refinamento dos modelos apresentados, dia após dia aparecem outras técnicas e novos modelos que são criados e testados na temática da bolsa de valores, como GRUB que é bastante similar ao GRU porém com arquitetura bidirecional ou o *Multilayer Perceptron* (MLP) que já é um modelo conhecido e está constantemente sendo explorado em novos cenários por exemplo.

Referências

- A.C, C. D. S. Redes neurais artificiais recorrentes aplicadas na correção de sinais distorcidos pela saturação de transformadores de corrente. In: *Sba: Controle Automação Sociedade Brasileira de Automatica*. [S.l.: s.n.], 2006. p. 424–436. Citado na página 21.
- AIRES, D. B.; DAMETTO, R. C.; CREPALDI, A. F. Previsão de séries temporais financeiras utilizando redes neurais artificiais: um comparativo na crise de 2008. São Paulo, Brazil, p. 30, 2017. Citado 2 vezes nas páginas 18 e 20.
- B3. *Ibovespa B3*. 2022. <b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm>. Citado 2 vezes nas páginas 18 e 19.
- BORELLA, L. et al. Aplicação de redes neurais artificiais e média móvel simples na previsão de demanda em uma empresa do setor metal-mecânico. Ponta Grossa, Brasil, p. 10, 2019. Citado na página 52.
- BORGES, L. E. *Python para desenvolvedores*. [S.l.]: novatec, 2014. Citado na página 35.
- BUTTON, S. T. Metodologia para planejamento experimental e análise de resultados. Campinas, Brasil, p. 88, 2021. Citado na página 44.
- CALVETTE, T. D. A. Aplicação de técnicas de machine learning e data driven em poços inteligentes de petróleo. Rio de Janeiro, Brazil, p. 68, 2019. Citado 2 vezes nas páginas 7 e 24.
- CAPES portal de periodicos. *Quem somos*. 2022. <<https://www-periodicos-capes-gov-br.ez1.periodicos.capes.gov.br/index.php/sobre/quem-somos.html>>. Citado na página 40.
- CHO KYUNGHYUN; VAN MERRIENBOER, B. G. C. B. D. B. F. S. H. B. Y. Learning phrase representations using rnn encoderdecoder for statistical machine translation. In: . [S.l.: s.n.], 2014. p. 1406. Citado na página 24.
- CHOPRA, R.; SHARMA, G. D. Application of artificial intelligence in stock market forecasting: A critique, review, and research agenda. Delhi, India, p. 23, 2021. Citado na página 14.
- CHUNG, H.; SHIN, K. shik. Genetic algorithm-optimized long short-term memory network for stock market prediction. Seoul, Coreia do Sul, p. 18, 2018. Citado 3 vezes nas páginas 7, 29 e 31.
- COLAH. *Understanding LSTM Networks*. 2015. <colah.github.io/posts/2015-08-Understanding-LSTMs/>. Citado 3 vezes nas páginas 7, 22 e 23.
- FERREIRA, A. S. Redes neurais convolucionais profundas na detecção de plantas daninhas em lavoura de soja. Mato Grosso do Sul, Brazil, p. 80, 2017. Citado na página 21.
- FLAUZINO, S. S. Redes neurais artificiais para engenharia e ciências aplicadas. In: *Artliber*. [S.l.: s.n.], 2010. p. 302. Citado 3 vezes nas páginas 7, 18 e 20.

- GERHARDT, T. E.; SILVEIRA, D. T. Métodos de pesquisa. Rio Grande do Sul, Brasil, p. 122, 2009. Citado 3 vezes nas páginas 38, 39 e 40.
- GIL, A. C. Como elaborar projetos de pesquisa. São Paulo, Brasil, p. 101, 1996. Citado 2 vezes nas páginas 38 e 39.
- GOMES, F. R. A bolsa de valores brasileira como fonte de informações financeiras. Belo Horizonte, Brazil, p. 14, 2007. Citado na página 17.
- INFOMONEY. *Mercado de Ações*. 2022. <infomoney.com.br/guias/mercado-de-acoes/>. Citado na página 18.
- J, P. et al. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. In: *Expert Syst. Appl.* [S.l.: s.n.], 2015. v. 1, p. 259–268. Citado na página 26.
- J.HOCHREITER. *Untersuchungen zu dynamischen neuronalen netzen*. Munchen, Germany, p. 74, 1991. Citado na página 22.
- JR, L.; COSTA, O.; E.M.M. *Redes neurais: fundamentos e aplicação com programas em c*. Rio de Janeiro, Brasil, 2007. Citado na página 19.
- JUNIOR, G. B. V. et al. Métricas utilizadas para avaliar a eficiência de classificadores em algoritmos inteligentes. Brasil, p. 14, 2022. Citado 3 vezes nas páginas 25, 27 e 28.
- JUPYTER. *Jupyter Notebook: The Classic Notebook Interface*. 2022. <<https://jupyter.org/>>. Citado na página 35.
- KALCHBRENNER, N.; GREFENSTETTE, E.; BLUNSON, P. A. A convolutional neural network for modelling senteces. p. 1404, 2014. Citado na página 21.
- KERAS. *Deep Learning for humans*. 2022. <<https://keras.io/>>. Citado na página 36.
- LAWI, A.; MESRA, H.; AMIR, S. Implementation of long short-term memory and gated recurrent units on grouped time-series data to predict stock prices accurately. Kyushu, Japão, p. 19, 2022. Citado 2 vezes nas páginas 29 e 31.
- LECUN, Y. B. Y.; HINTON, G. Deep learning. 2015. Citado na página 19.
- LV, P. et al. Stock index prediction based on time series decomposition and hybrid model. Chengdu, China, p. 18, 2022. Citado 3 vezes nas páginas 31, 62 e 63.
- MARTINS, M. A. dos S.; METTE, F.; MACEDO, G. R. de. A utilização de redes neurais artificiais para a estimativa dos preços da petrobrás pn na bovespa. Rio Grande do Sul, Brasil, p. 16, 2008. Citado na página 14.
- M.A.S, M.; F, M.; MACEDO, G. A utilização de redes neurais artificiais para a estimação dos preços da petrobrás pn na bovespa. p. 16, 2008. Citado na página 19.
- MATPLOTLIB. *Matplotlib: Visualization with Python*. 2022. <<https://matplotlib.org/>>. Citado na página 36.
- MONTGOMERY, D. C. *Montgomery's design and analysis of experiments*. Estados Unidos, p. 730, 1991. Citado na página 45.

- NABIPOUR, M. et al. Deep learning for stock market prediction. Teerão, Irã, p. 268, 2020. Citado 5 vezes nas páginas 20, 22, 26, 28 e 31.
- NABIPOUR, M. et al. Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. Iran, p. 14, 2020. Citado na página 15.
- NAPOLEÃO, B. M. Estabelecendo uma string de busca para a identificação de estudos secundários na engenharia de software. Paraná, Brasil, p. 54, 2019. Citado na página 40.
- NETO, G. C. M. O desempenho do método de análise técnica bandas bollinger. Uberlândia, Brazil, p. 20, 2018. Citado na página 53.
- NUMPY. *Numpy, get started*. 2022. <<https://numpy.org/>>. Citado na página 36.
- P. BA, J. L. K. D. Adam: a method for stochastic optimization. 2019. Citado na página 57.
- PAGE, M. J. Declaración prisma 2020: una guía actualizada para la publicación de revisiones sistemáticas. Espanha, p. 10, 2021. Citado na página 41.
- PANDAS. *Pandas*. 2022. <<https://pandas.pydata.org/>>. Citado na página 35.
- PASQUOTTO, J. L. D. Previsão de séries temporais no varejo brasileiro: Uma investigação comparativa da aplicação de redes neurais recorrentes de elman. São Paulo, Brasil, p. 200, 2010. Citado 2 vezes nas páginas 56 e 57.
- PAZ, R. R. C. da. Utilizando um algoritmo lstm para previsão do preço de uma ação. Campina Grande, Brazil, p. 12, 2019. Citado na página 56.
- PENNA, M. L. F. Rede neural artificial para detecção de sobremortalidade atribuível a colera no ceará. In: *Revista de Saude Publica*. [S.l.: s.n.], 2004. p. 351–357. Citado na página 22.
- SAKO BERTHINE NYUNGA MPINDA, P. C. R. K. Neural networks for financial time series forecasting. Basel, Switzerland, p. 17, 2022. Citado na página 51.
- SANVICENTE, H. Índice bovespa: um padrão para os investimentos brasileiros. São Paulo, Brasil, p. 140, 1994. Citado na página 18.
- SAS. *Deep Learning O que é e qual sua importância?* 2022. <sas.com/pt_br/insights/analytics/deep-learning.html>. Citado na página 19.
- SCM git. *Git*. 2022. <<https://git-scm.com/>>. Citado na página 33.
- SOUZA, A. de. Bovespa: Uma análise do desempenho no período pós-real. p. 71, 2008. Citado na página 19.
- SOUZA, V. et al. Análise comparativa de redes neurais convolucionais no reconhecimento de cenas. Marabá, Brazil, p. 8, 2020. Citado na página 21.
- TEAM, E. I. D. *Oracle Cloud Como Trabalhar com o Planning*. [S.l.]: Oracle, 2022. Citado na página 26.

VARGAS, A. C. G.; CARVALHO, A. M. P.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. Niteroi, Brazil, p. 4, 2017. Citado 2 vezes nas páginas 7 e 21.

WALT, V. der et al. Scikit-image: image processing in python. p. 453, 2014. Citado na página 21.

WANG, W.-J. et al. Stock market index prediction based on reservoir computing models. Guangxi, China, p. 18, 2022. Citado 4 vezes nas páginas 29, 31, 62 e 63.

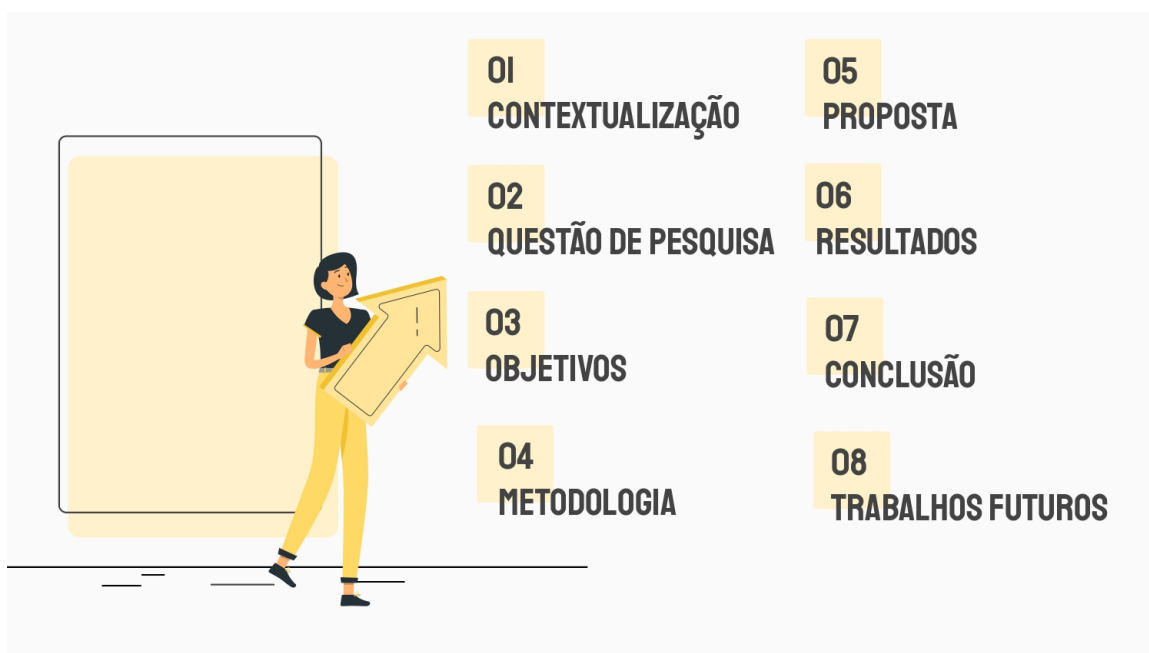
WYTHOFF, B. J. Backpropagation neural networks: A tutorial. *Chemometrics and Intelligent Laboratory Systems*, v. 18, n. 2, p. 115–155, 1993. ISSN 0169-7439. Disponível em: <<https://www.sciencedirect.com/science/article/pii/016974399380052J>>. Citado 2 vezes nas páginas 7 e 44.

YAHOO. Citado 2 vezes nas páginas 7 e 50.

ZHENG, H.; WANG, H.; CHEN, J. Evolutionary framework with bidirectional long short-term memory network for stock price prediction. Shenzhen, p. 8, 2020. Citado na página 51.

.1 Slides de apresentação do trabalho final

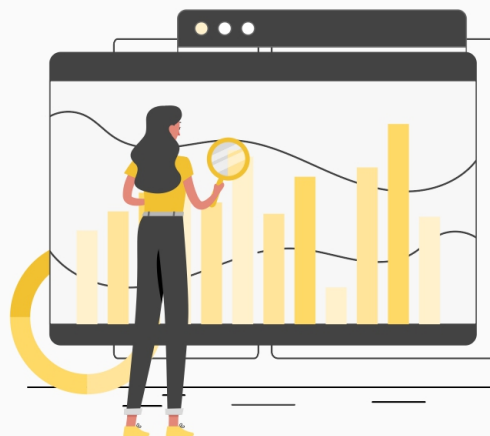




01.

CONTEXTUALIZAÇÃO

Uma descrição da conjuntura do problema trabalhado



CONTEXTUALIZAÇÃO



CHOPRA; SHARMA; 2021

A previsão de ações concentra-se em estimar o movimento futuro do preço de uma ação, que é geralmente percebida como uma tarefa desafiadora devido ao ruído e características voláteis.



MARTINS; METTE; MACEDO; 2008

A utilização de técnicas de inteligência artificial para prever preços de ativos tem demonstrado resultados positivos, em que as redes neurais artificiais podem tratar o tema de maneira mais simples que os métodos não lineares.

CONTEXTUALIZAÇÃO

**FREIRE; MENEZES; BARRETO, 2009**

Redes neurais artificiais (RNAs) têm sido utilizadas com sucesso em tarefas de predição e modelagem de séries temporais

**HOCHREITER; SCHMIDHUBER, 1997**

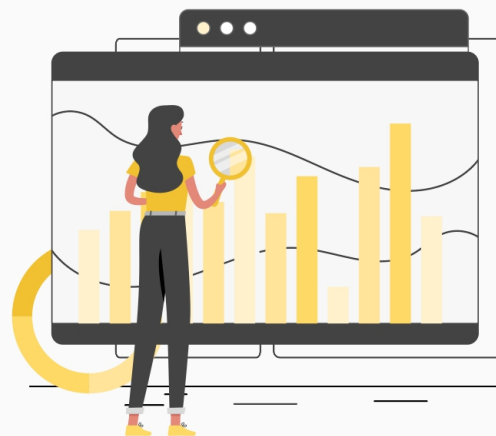
O LSTM leva a muito mais execuções bem-sucedidas e aprende muito mais rápido.

**CHO KYUNGHYUN, VAN MERREINBOER**

A rede neural GRU é uma versão atualizada e compacta da rede LSTM. Demanda menos *gates* e parâmetros

02. QUESTÃO DE PESQUISA

O que esse trabalho procura responder?



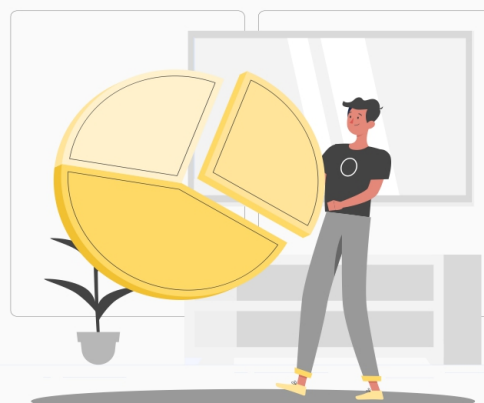


É POSSÍVEL UTILIZAR REDES NEURAIS RECORRENTES PARA PREVISIONAR O ÍNDICE DA BOLSA DE VALORES BRASILEIRA?

03.

OBJETIVOS

Apresentação do objetivo geral e principais objetivos específicos do trabalho



OBJETIVO GERAL

- Estudar a viabilidade do uso de redes neurais recorrentes na predição do índice da bolsa de valores brasileira, realizando um comparativo entre as arquiteturas LSTM e GRU, observando diferentes horizontes de tempo.

OBJETIVOS ESPECÍFICOS**1**

Avaliar dados históricos e verificar o efeito das técnicas de pré-processamento de dados;

2

Estudo e definição dos melhores hiperparâmetros;

3

Avaliar a qualidade das previsões em um espaço de tempo de 1, 5, 30 e 90 dias;

4

Comparativo técnico entre resultados obtidos.

04.

METODOLOGIA

Pesquisa, Desenvolvimento e Análise



04.

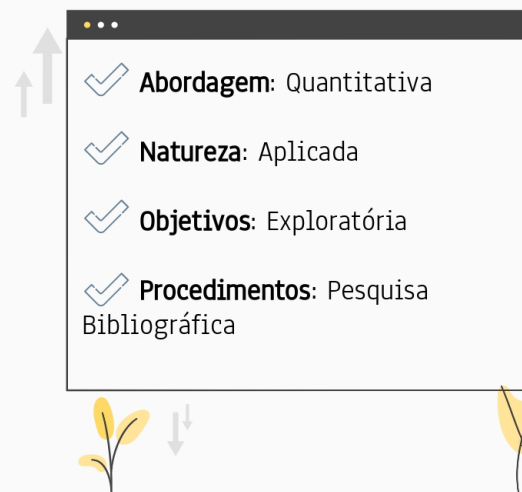
METODOLOGIA

Pesquisa, Desenvolvimento e Análise



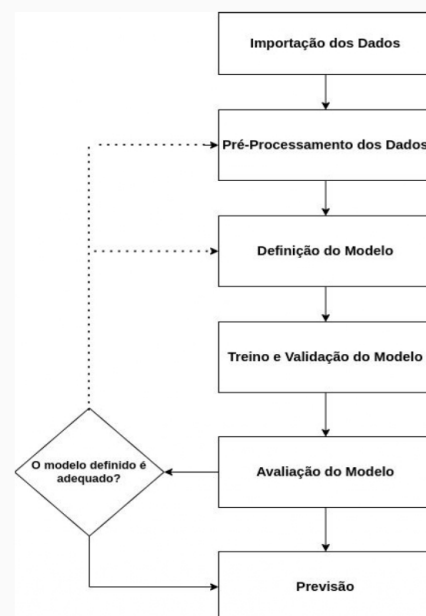
METODOLOGIA DE PESQUISA

- *"Pesquisa Científica é o resultado do inquérito ou exame minucioso realizado com o objetivo de resolver um problema, recorrendo a procedimentos científicos."* (GERHARDT; SILVEIRA, 2009)



METODOLOGIA DE DESENVOLVIMENTO

Metodologia adotada para desenvolvimento dos algoritmos de previsão.





05.

PROPOSTA

O que foi proposto e desenvolvido

PROPOSTA

ARQUITETURA DE PREDIÇÃO: GRU

Desenvolvimento de uma rede neural recorrente GRU para predição do índice do ibovespa representado pela ação BOVA11 nos horizontes de tempo: 1, 5, 30, 90 dias.

ARQUITETURA DE PREDIÇÃO: LSTM

Desenvolvimento de uma rede neural recorrente LSTM para predição do índice do bovespa representado pela ação BOVA11 nos horizontes de tempo: 1, 5, 30, 90 dias.

COLETA DOS DADOS

Dados de 2 de Janeiro de 2009 a 02 de Fevereiro de 2023, totalizando 2497 registros com informações de:

- Data de negociação;
- Valor de Abertura;
- Valor máximo;
- Valor mínimo;
- Valor de fechamento e
- Volume

	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
0	2009-01-02 00:00:00-02:00	38.279999	40.320000	37.939999	40.320000	119900	0	0
1	2009-01-05 00:00:00-02:00	40.200001	42.000000	39.700001	42.000000	341700	0	0
2	2009-01-06 00:00:00-02:00	42.060001	42.320000	41.549999	42.320000	178300	0	0
3	2009-01-07 00:00:00-02:00	41.750000	41.770000	40.450001	40.639999	73400	0	0
4	2009-01-08 00:00:00-02:00	40.150002	42.200001	40.150002	42.200001	195300	0	0
5	2009-01-09 00:00:00-02:00	41.980000	42.470001	41.160000	42.000000	271900	0	0
6	2009-01-12 00:00:00-02:00	40.709999	40.750000	39.380001	39.700001	37700	0	0
7	2009-01-13 00:00:00-02:00	0.000000	40.259998	38.540001	40.000000	177300	0	0
8	2009-01-14 00:00:00-02:00	39.400002	39.470001	37.700001	38.029999	205500	0	0
9	2009-01-15 00:00:00-02:00	37.889999	39.360001	36.930000	39.360001	120500	0	0

PRÉ-PROCESSAMENTO

- Remover linhas de volume 0;
- Ordenar o dataframe pela data;
- Adicionar colunas de médias móveis de 3, 10, 30 e 60 dias;
- Adicionar Colunas de desvio padrão de 7, 30 e 60 dias;
- Adicionar Bandas de Bollinger superiores e inferiores de 3, 30 e 60 dias;
- Adicionar a força relativa no período de 15 dias
- Normalizar os dados numa escala de -1 a 1

```
[[-0.41052802 -0.95937954 -0.97594966 ... -0.34158244 -0.812735  
-0.94288915]  
[-0.39380423 -0.93549828 -0.95100022 ... -0.1145866 -0.81635782  
-0.94311308]  
[-0.38663696 -0.9468809 -0.94178467 ... -0.11645738 -0.82917641  
-0.90033593]  
...  
[ 0.76315995 0.6625376 0.66621716 ... 0.35221771 -0.85035523  
0.6114222 ]  
[ 0.75647048 0.64445932 0.63205208 ... 0.16001159 -0.89494206  
0.61814103]  
[ 0.73624277 0.62280992 0.62733194 ... 0.1831316 -0.85119142  
0.64367293]]
```

MODELAGEM

Construção dos modelos LSTM e GRU utilizando a interface de implementação TensorFlow.



CONSTRUÇÃO DAS REDES NEURAIS

- Sequential - Para inicializar a rede neural;
- LSTM / GRU para adicionar a camada de memória;
- Dropout para ajuste entre camadas;
- Dense - Para adicionar uma camada de "n" predições densamente conectada;
- Early Stopping para parada antecipada do treinamento;

HYPERPARÂMETRO TUNING

- Número de neurônios: 64, 128 ou 256;
- Batch Size: 32 ou 64;
- Dropout: 0.2 ou 0.1;
- Janela de tempo: 50 ou 100 dias passados;

Para validação da melhor combinação de hiperparâmetros foi utilizado a métrica de erro RMSE

Tunagem de hiperparâmetros da rede GRU

Unidade	Batch Size	Dropout	Janela de tempo	RMSE
256	32	0.1	100	0.0235
256	32	0.2	50	0.0241
256	64	0.1	50	0.0247
256	64	0.1	100	0.0247
128	32	0.2	100	0.0252

Tunagem de hiperparâmetros da rede LSTM

Unidade	Batch Size	Dropout	Janela de Tempo	RMSE
256	32	0.1	50	0,0224
256	32	0.2	100	0,0264
64	64	0.1	100	0,0276
128	32	0.1	100	0,0280
256	64	0.2	100	0,0282

06.

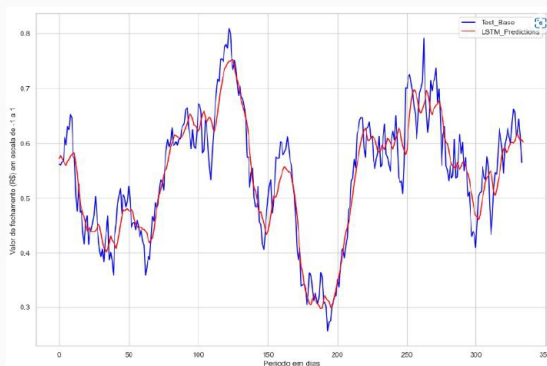
RESULTADOS E DISCUSSÕES

Resultados e análises acerca do trabalho desenvolvido

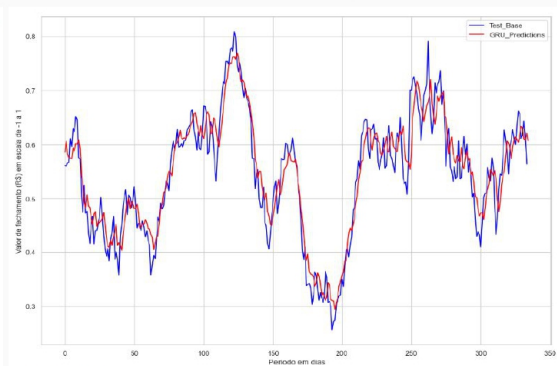


RESULTADO DOS TESTES- LSTM E GRU

Predição da rede LSTM



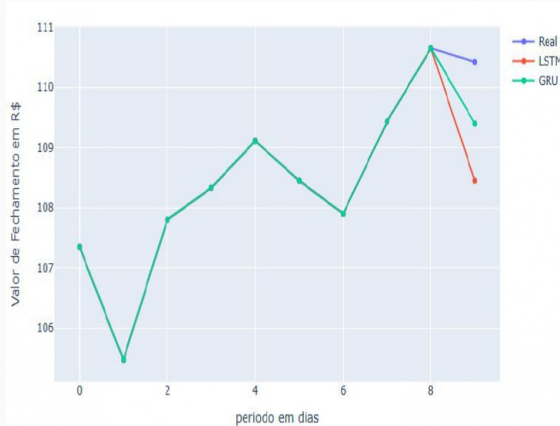
Predição da rede GRU



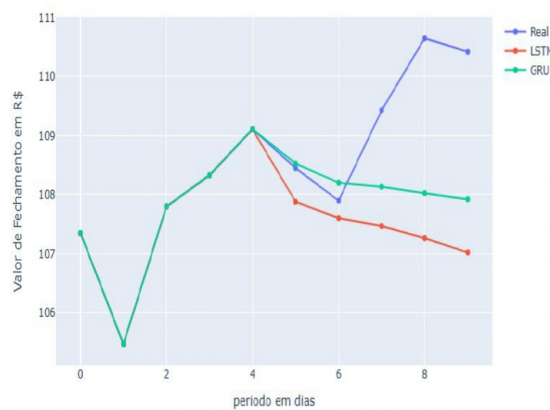
Rede	MAE	MAPE	RMSE
LSTM	0.032	6.11	0.040
GRU	0.031	6.10	0.038

PREVENDO 1 E 5 DIAS A FRENTE

Previsão de 1 dia

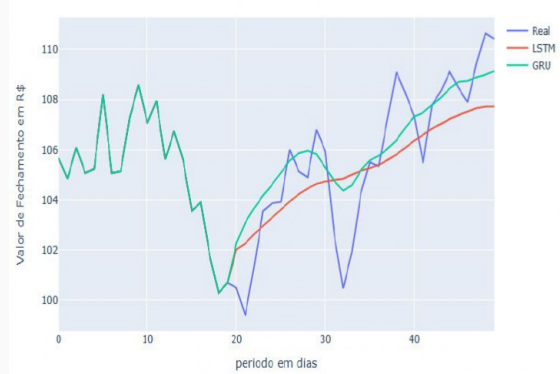


Previsão de 5 dias

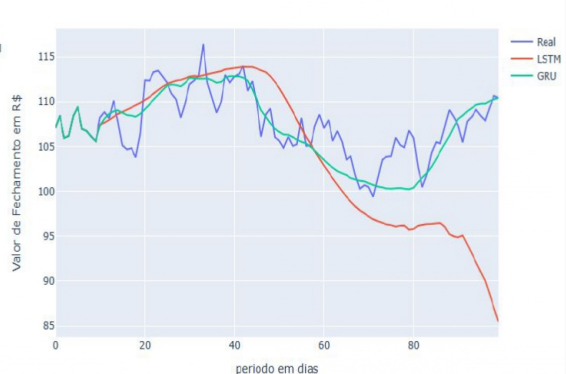


PREVENDO 30 E 90 DIAS A FRENTE

Previsão de 30 dias



Previsão de 90 dias

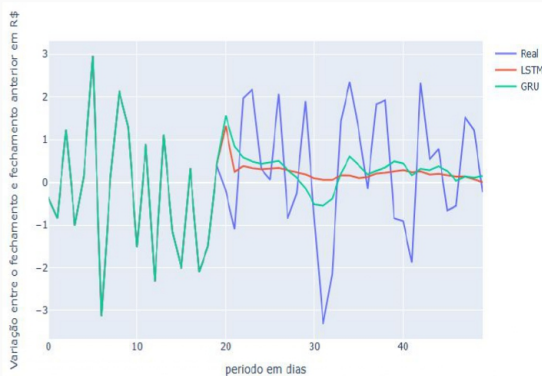


PREVENDO 30 E 90 DIAS A FRENTE

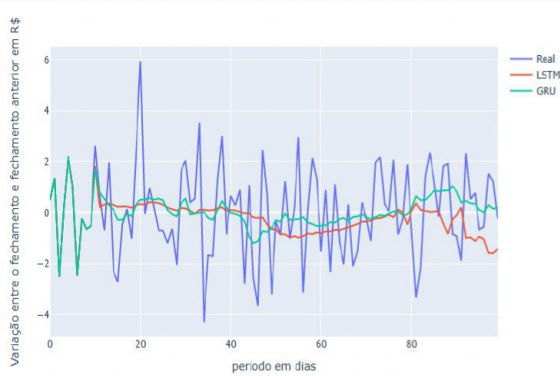
Rede	Dias de Predição	MAE	MAPE	RMSE
LSTM	1	0,048	7,86	0,048
GRU	1	0,069	10,95	0,069
LSTM	5	0,026	4,4	0,030
GRU	5	0,019	3,21	0,024
LSTM	30	0,047	8,92	0,053
GRU	30	0,041	7,50	0,051
LSTM	90	0,13	44,74	0,18
GRU	90	0,056	10,17	0,068

ENTENDENDO A VARIAÇÃO DIA A DIA DAS REDES NEURAIAS

Varição em relação ao dia anterior na predição de 30 dias



Varição em relação ao dia anterior na predição de 90 dias



07. CONCLUSÕES

Atualização dos status dos objetivos e atividades propostas



STATUS DOS OBJETIVOS

1

Avaliar dados históricos e verificar o efeito qualitativo das técnicas de pré-processamento de dados;
Feito

2

Estudo e definição dos melhores hiperparâmetros;
Feito

3

Avaliar a qualidade das previsões em um espaço de tempo de 1, 5, 30 e 90 dias;
Feito

4

Comparativo técnico entre resultados obtidos
Feito

STATUS DO TRABALHO DE CONCLUSÃO DE CURSO 2

Atividades	Status
Aplicação das correções recomendadas	Concluída
Desenvolvimento dos algoritmos	Concluída
Análise dos resultados	Concluída
Descrição dos resultados	Concluída
Revisão e aplicação das correções finais	Concluída
Apresentação do TCC 2	Em Andamento

08. TRABALHOS FUTUROS

O que pode ser melhorado, refinado ou estudado de forma diferente?



TRABALHOS FUTUROS

**01
PRÉ-PROCESSAMENTO**

O pré-processamento é um processo extremamente complexo, podendo ter diversos pontos de melhoria como identificação de *outliers*, sazonalidades, novas features, entre outras melhorias.

**03
TESTAR NOVOS MODELOS**

No trabalho realizado, foi feito um comparativo entre LSTM e GRU, outros modelos como GRUB e MLP podem apresentar resultados interessantes para estudo.

**02
MODELAGEM**

A modelagem apresentada, testou algumas combinações de hiperparâmetros. Pode-se realizar testes com outras combinações ou testar outros hiperparâmetros a fim de se obter um resultado melhor.

**04
TESTE COM OUTRAS ENTRADAS**

O trabalho tentou realizar a predição do índice do Bovespa, que conta com altas variações diárias. As redes recursivas podem ser testadas na predição, por exemplo dos FII's, que apresentam uma variação menor.

OBRIGADO!