

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

CAIO NAIM PAES FERREIRA
GABRIEL RIBAS PENHA DAS CHAGAS
LUCAS DA SILVA FARIAS

USO DE SÉRIES TEMPORAIS PARA ANÁLISE QUANTITATIVA DE AÇÕES EM
MEIO À COVID-19

RIO DE JANEIRO
2023

CAIO NAIM PAES FERREIRA
GABRIEL RIBAS PENHA DAS CHAGAS
LUCAS DA SILVA FARIAS

USO DE SÉRIES TEMPORAIS PARA ANÁLISE QUANTITATIVA DE AÇÕES EM
MEIO À COVID-19

Trabalho de conclusão de curso de graduação
apresentado ao Instituto de Computação da
Universidade Federal do Rio de Janeiro como
parte dos requisitos para obtenção do grau de
Bacharel em Ciência da Computação.

Orientador: Valeria Menezes Bastos

RIO DE JANEIRO
2023

F383u

Ferreira, Caio Naim Paes

Uso de séries temporais para análise quantitativa de ações em meio à Covid-19 / Caio Naim Paes Ferreira, Gabriel Ribas Penha das Chagas e Lucas da Silva Farias. – 2023.

61 f.

Orientadora: Valeria Menezes Bastos.

Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação)
- Universidade Federal do Rio de Janeiro, Instituto de Computação, Bacharel em Ciência da Computação, 2023.

1. Inteligência artificial. 2. Aprendizado de máquina. 3. Redes neurais. 4. Séries temporais. 5. Mercado de ações. 6. Ações. I. Chagas, Gabriel Ribas Penha das. II. Farias, Lucas da Silva. III. Bastos, Valeria Menezes (Orient.). IV. Universidade Federal do Rio de Janeiro, Instituto de Computação. V. Título.

CAIO NAIM PAES FERREIRA
GABRIEL RIBAS PENHA DAS CHAGAS
LUCAS DA SILVA FARIAS

USO DE SÉRIES TEMPORAIS PARA ANÁLISE QUANTITATIVA DE AÇÕES EM
MEIO À COVID-19

Trabalho de conclusão de curso de graduação
apresentado ao Instituto de Computação da
Universidade Federal do Rio de Janeiro como
parte dos requisitos para obtenção do grau de
Bacharel em Ciência da Computação.

Aprovado em 14 de junho de 2023

BANCA EXAMINADORA:

Valeria Menezes Bastos
D.Sc. (UFRJ)

Documento assinado digitalmente
 DANIEL SADOC MENASCHE
Data: 03/08/2023 21:35:43-0300
Verifique em <https://validar.iti.gov.br>

Daniel Sadoc Menasché
Ph.D. (UMass)

Documento assinado digitalmente
 CLAUDIO MICELI DE FARIAS
Data: 08/08/2023 15:06:41-0300
Verifique em <https://validar.iti.gov.br>

Claudio Miceli de Farias
D.Sc. (UFRJ)

Eu, Gabriel Ribas Penha das Chagas, gostaria de dedicar este trabalho ao meu avô, Jorge Ribas e ao meu pai, Jorge Penha. Infelizmente não estão comigo hoje, mas foram minha base durante o processo. À minha mãe, Hilda Ribas, e às minhas irmãs, Raquel e Vitória, por serem o motivo de eu acordar todo dia para seguir em frente. Aos meus amigos, Caio Fabio, Rodrigo Lopes, Igor Guedes, Caio Paranhos, Mateus Dantas e Erick Mota, irmãos que a vida me deu. Agradecer a Julia, por ter me apoiado e me aconselhado durante o desenvolvimento desse trabalho.

Eu, Caio Naim Paes Ferreira, gostaria de dedicar este trabalho ao meu pai Robson Paes Ferreira, minha mãe, Alexandra Naim Paes Ferreira e minha tia Isa Maria Paes ferreira. Sem dúvidas, os 3 maiores pilares da minha educação que, com muita dedicação, me permitiram chegar até onde estou agora. Também dedico à minha irmã, Camilla Naim Paes Ferreira, por sempre se preocupar com minha saúde mental e me motivar à sempre procurar a felicidade e o amor. Por fim, dedico a todos os meus amigos que fizeram, e fazem ainda, parte da minha jornada e caminham ao meu lado e à minha namorada, Rebecca de Araujo Coutinho, por me sustentar e motivar independente das adversidades.

Eu, Lucas da Silva Farias, gostaria de dedicar este trabalho à minha mãe, Jocélia Pedro da Silva Farias e ao meu pai, Erisvaldo Farias Ribeiro. São os responsáveis por me apoiar em todas as decisões e permitir que chegasse até aqui, além de sempre proporcionarem as melhores condições para que eu seguisse em frente com os estudos. Também deixo dedicado aos meus amigos, que participaram de todo esse processo, por mais breve que tenha sido, e em especial à minha namorada Mariana Glória e minha amiga Pryscila Santiago, por serem quem mais me ouviram, aconselharam e apoiaram durante todos esses anos de graduação.

AGRADECIMENTOS

Gostaríamos de agradecer a instituição UFRJ, ao departamento de Ciência da Computação e a todos os professores que fizeram parte dessa jornada de aprendizado. Em especial, à professora Valeria Menezes Bastos, por ser além de uma excepcional professora e educadora, uma grande amiga.

*“I never dreamed about success.
I worked for it.”*

Estée Lauder

RESUMO

O estudo foi realizado buscando entender a consistência de modelos preditivos frente a impactos externos que influenciam diretamente a economia, como, por exemplo, a COVID-19. A análise consistiu em avaliar a acurácia de 3 abordagens: Simulação de Monte Carlo, Séries Temporais e Long Short-Term Memory (LSTM) antes do acontecimento da pandemia. Avaliaram-se as previsões para o cenário de COVID-19 e foi feito um comparativo com o real, a fim de avaliar a robustez de modelos em meio ao cenário econômico mundial. O resultado obtido mostrou que apenas a variável de tempo não é o suficiente para lidar com previsão de preços no mercado de ações, tornando até mesmo o modelo mais robusto suscetível a falhas nas previsões em cenários pouco incertos. Portanto, há a necessidade de avaliar, juntamente aos padrões de evolução no tempo, fatores como notícias políticas, financeiras, redes sociais e diversas outras fontes que possam correlacionar com a variável de tempo e prover previsões mais assertivas para cenários completos.

Palavras-chave: inteligência artificial; aprendizado de máquina; redes neurais; séries temporais; mercado de ações; ações.

ABSTRACT

The study was carried out seeking to understand the consistency of predictive models in the face of external impacts that directly influence the economy, such as COVID-19. The analysis consisted of evaluating the accuracy of 3 approaches: Monte Carlo Simulation, Time Series and Long Short-Term Memory (LSTM) before the pandemic event. The predictions for the COVID-19 scenario were evaluated, and a comparison was made with the real one, in order to evaluate the robustness of models in the midst of the world economic scenario. The result obtained showed that the time variable alone is not enough to deal with price prediction in the stock market, making even the most robust model susceptible to failure in predictions in less uncertain scenarios. Therefore, there is a need to evaluate, together with the patterns of evolution over time, factors such as political, financial news, social networks and several other sources that can correlate with the time variable and provide more assertive predictions for complete scenarios.

Keywords: artificial intelligence; machine learning; neural network; time series; stock market; stocks;

LISTA DE ILUSTRAÇÕES

Figura 1 – Pessoas Físicas na Bolsa 2011 – 2021	12
Figura 2 – Interesse em ações ao longo do tempo	13
Figura 3 – Meios utilizados para efetuar o investimento.	14
Figura 4 – Exemplo de Distribuição Normal	18
Figura 5 – Venda de material de construção em função do tempo	19
Figura 6 – Venda de material de construção em função do tempo	21
Figura 7 – Série não estacionária	23
Figura 8 – Valores antes e depois da diferenciação	23
Figura 9 – Gráfico após diferenciação	24
Figura 10 – gráficos de autocorrelação e autocorrelação parcial	25
Figura 11 – Estrutura de um neurônio	26
Figura 12 – Modelo de um neurônio artificial	27
Figura 13 – Tipos de Redes Neurais	28
Figura 14 – Gráfico da Função Sigmoide	29
Figura 15 – Gráfico da Função Tangente Hiperbólica	30
Figura 16 – Exemplo de arquitetura de uma RNA com duas camadas ocultas	31
Figura 17 – Arquitetura simples de uma RNN com um único neurônio e com três	33
Figura 18 – Comparação entre células RNN e LSTM	34
Figura 19 – Estado da célula — LSTM	35
Figura 20 – Forget Gate — LSTM	35
Figura 21 – Input Gate — LSTM	36
Figura 22 – Atualização do estado da célula	37
Figura 23 – Output Gate — LSTM	37
Figura 24 – Preços de fechamentos da ação GOLL4	40
Figura 25 – Preços de fechamentos da ação PETR4	41
Figura 26 – Preços de fechamentos da ação MGLU3	42
Figura 27 – Previsões GOL — Cenário 1	45
Figura 28 – Previsões GOL — Cenário 2	46
Figura 29 – Previsões GOL — Cenário 3	46
Figura 30 – Previsões Petrobras — Cenário 1	47
Figura 31 – Previsões Petrobras — Cenário 2	47
Figura 32 – Previsões Petrobras — Cenário 3	47
Figura 33 – Magalu — Cenário 1	48
Figura 34 – Magalu — Cenário 2	48
Figura 35 – Magalu — Cenário 3	49
Figura 36 – Previsões GOL Modelo Sarima	50

Figura 37 – Previsões Petrobras Modelo Sarima	51
Figura 38 – Previsões Magazine Luiza Modelo Sarima	52
Figura 39 – Previsões GOL para o LSTM	53
Figura 40 – Previsões Petrobras para o LSTM	54
Figura 41 – Previsões Magalu para o LSTM	56

SUMÁRIO

1	INTRODUÇÃO	12
1.1	OBJETIVO	15
1.1.1	Geral	15
1.1.2	Específicos	15
1.2	ESCOPO DO TRABALHO	15
1.3	ESTRUTURA DO TRABALHO	15
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	SIMULAÇÃO DE MONTE CARLO	17
2.2	SÉRIES TEMPORAIS	18
2.2.1	Modelo SARIMA e antecessores	19
2.2.1.1	Fundamentos teóricos	20
2.3	REDES NEURAIS ARTIFICIAIS	25
2.3.1	Função de Ativação	27
2.3.2	Arquitetura	29
2.3.3	Processo de Aprendizado	30
2.3.3.1	Aprendizado Supervisionado	32
2.3.3.2	Aprendizado Não Supervisionado	32
2.3.4	Rede Neural Recorrente	32
2.3.5	Rede Long Short-Term Memory	34
3	TRABALHOS RELACIONADOS	38
4	DESENVOLVIMENTO DA PROPOSTA	39
4.1	DADOS	39
4.2	CENÁRIOS ANALISADOS	39
4.3	TRATAMENTO DOS DADOS	40
4.4	CONSTRUÇÃO DOS MODELOS	41
4.4.1	Monte Carlo	41
4.4.2	Sarima	41
4.4.3	LSTM	42
5	ANÁLISE DOS RESULTADOS	45
5.1	MONTE CARLO	45
5.1.1	GOL	45
5.1.2	Petrobras	46

5.1.3	Magazine Luiza	48
5.2	SARIMA	49
5.2.1	GOL	49
5.2.2	Petrobras	49
5.2.3	Magalu	50
5.3	REDES NEURAS	52
5.3.1	GOL	53
5.3.2	Petrobras	54
5.3.3	Magazine Luiza	55
5.3.4	Análise das previsões	55
6	CONCLUSÃO	57
6.1	CONCLUSÃO DOS MODELOS DE PREVISÃO PARA OS ATIVOS E PERÍODOS ESCOLHIDOS	57
6.1.1	Conclusão Monte Carlo	57
6.1.2	Conclusão SARIMA	57
6.1.3	Conclusão LSTM	58
6.2	CONCLUSÃO COMPARANDO OS 3 MODELOS	58
6.3	CONCLUSÃO SOBRE DESENVOLVIMENTO EM GRUPO	58
	REFERÊNCIAS	60

1 INTRODUÇÃO

Neste capítulo, será apresentado uma visão geral sobre o tema abordado e o porquê de analisar esse cenário, bem como os objetivos por trás do trabalho.

Em dezembro de 2021, o número de investidores pessoas físicas (PFs) na Bolsa brasileira (B3) atingiu 4.976.145, como pode ser visto na Figura 1. Em relação ao mês anterior, houve um aumento de 888.347 investidores PFs, equivalente a um crescimento mensal impressionante de +27,1% — o maior percentual de crescimento de 2021. Comparado a 2020, quando encerramos o ano com 3.229.318 contas ativas na B3, o ano fechou com uma alta de +54,1%. Acreditamos que esse crescimento de longo prazo está alinhado com a melhora da educação financeira no país.

Figura 1 – Pessoas Físicas na Bolsa 2011 – 2021



Fonte: B3, XP Investimentos

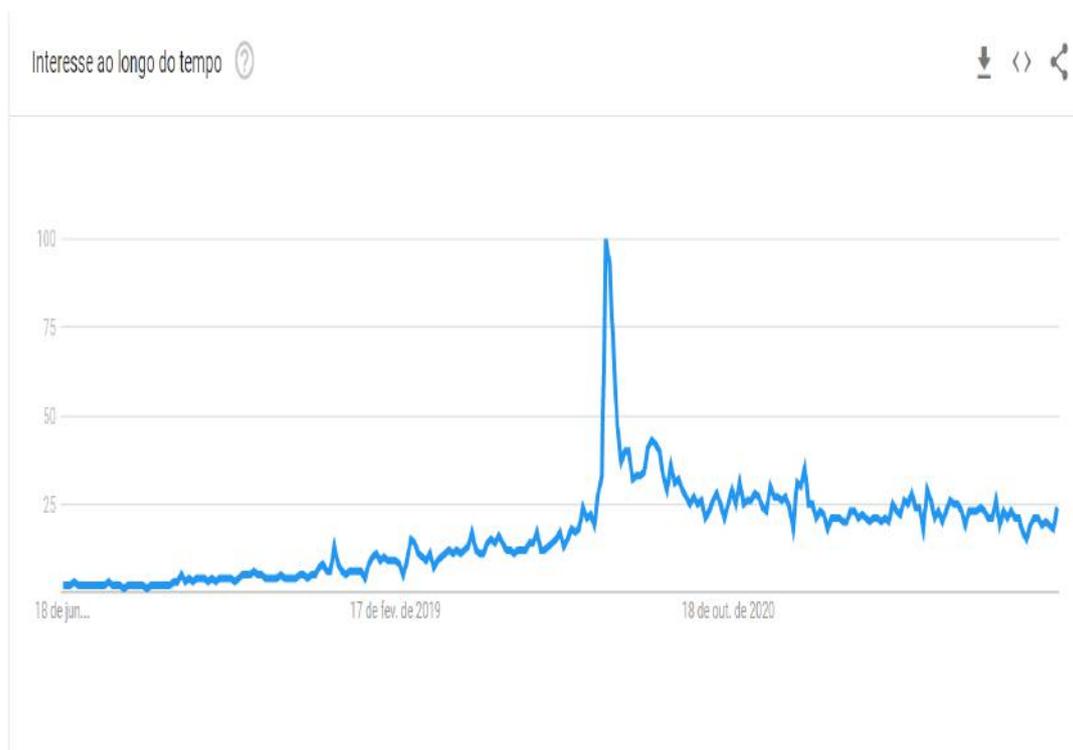


Fonte: XP Investimentos

Nos últimos cinco anos, a análise de ações e índices elevam o interesse de cada vez mais pessoas. Este fato pode ser evidenciado no contexto brasileiro, pela quantidade de pesquisas realizadas no site da Google, conforme mostrado na Figura 2, com a descrição “IBOVESPA”, o principal índice brasileiro.

Os meios digitais ganharam espaço na hora de fazer investimento. Em 2020, pela primeira vez, o aplicativo do banco foi a solução mais utilizada, ultrapassando a ida presencial à instituição, que liderava nos dois anos anteriores. O uso do aplicativo mais

Figura 2 – Interesse em ações ao longo do tempo



Fonte: Google trends pesquisando por "IBOVESPA"

do que dobrou em 2020, passando de 30% para 62% na preferência como meio para uma aplicação financeira. A visita presencial à agência caiu de 71% em 2019 para 55% um ano depois.

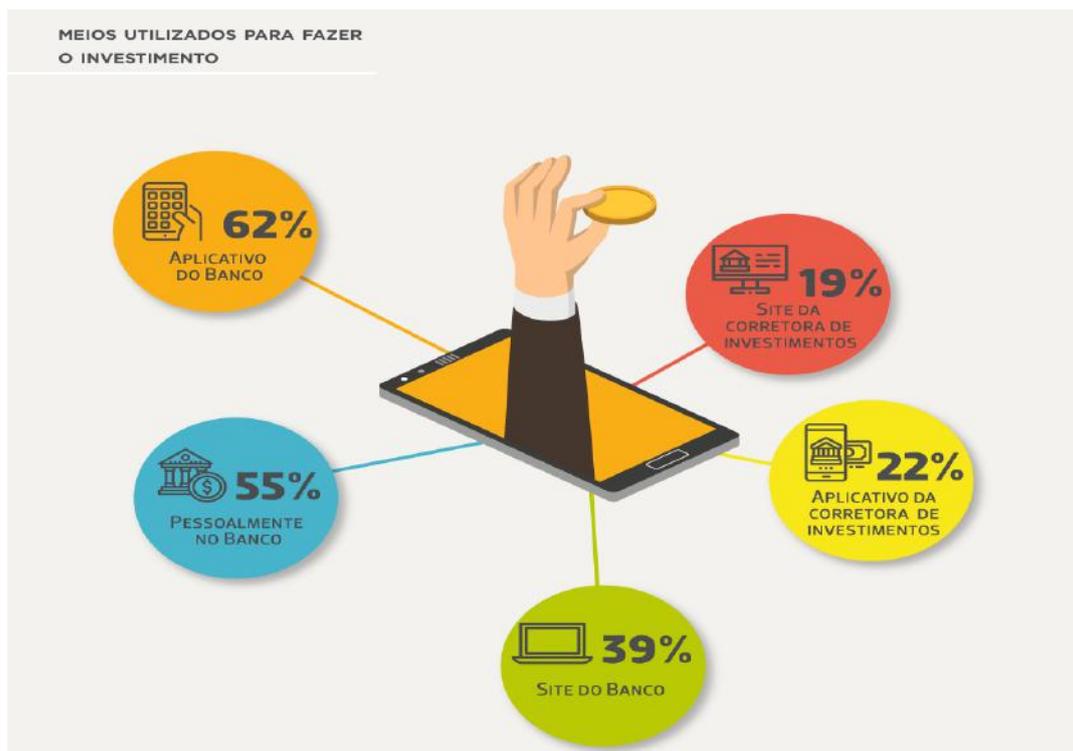
Todas as outras formas de investimento a distância também apresentaram crescimento significativo, como site do banco ou corretora e por telefone.

Correlacionando a Figura 2 com a Figura 3, notamos que o aumento da digitalização e o interesse por investimentos ganharam força em uma época próxima, no início da pandemia devido à COVID-19. Em meio a estes avanços, tornou-se factível utilizar modelos computacionais para auxiliar a tomada de decisão em meio a falta de especialização financeira.

Prever movimentações nos mercados financeiros é uma tarefa complicada devido às incertezas envolvidas. Segundo Patel et al., (2015), existem dois tipos de análises que os investidores podem realizar antes de investir em uma determinada ação. O primeiro tipo é a análise fundamentalista. Nesta modalidade, os investidores analisam o valor intrínseco das ações, considerando fatores como o desempenho de produção, as vendas, a saúde econômica da empresa e clima político, entre outros pontos para decidir sobre a compra ou venda do ativo. Outro tipo de análise é a técnica. Neste caso, a avaliação das ações considera o estudo da variação do preço e volume ao longo do tempo.

Contudo, prever valores resultantes no tempo não é uma tarefa simples, ainda mais

Figura 3 – Meios utilizados para efetuar o investimento.



Fonte: (ANBIMA, 2021)

no contexto de mercados financeiros. A razão se deve às séries temporais financeiras que por natureza são estocásticas, ruidosas, e não estacionárias, como explicado por Manav Kaushik et al., (2020). Estas características corroboram a Hipótese de Mercados Adaptativos (HMA) proposta por Lo (2004), que trata da conciliação da Hipótese de Eficiência de Mercados (HEM) com a economia comportamental, onde a oscilação de valor de uma ação pode ser afetada não somente pelo comportamento do histórico do preço mas também por informações macroeconômicas e sentimentos de investidores, tais como aversão à perda e excesso de confiança. Sabendo disso, é possível observar ao longo do tempo que determinados padrões de tomada de decisão dos investidores tornam-se recorrentes, isso faz com que sejam criadas várias estratégias para maximizar os lucros em negociações. Segundo Lo (2004), como nem toda informação financeira é explícita no mercado, ineficiências acabam afetando as estratégias de negociação.

Complementando o aumento do interesse no mercado financeiro, estamos atualmente lidando com as consequências de uma pandemia global que se estendeu por mais de um ano. É inegável o impacto significativo que a pandemia de COVID-19 teve sobre todas as empresas listadas na bolsa de valores. Diante desse contexto, surge a motivação de explorar a utilização da pandemia como um teste para avaliar a capacidade dos modelos preditivos em se adaptarem a situações excepcionais e identificar as discrepâncias em relação às previsões apresentadas em numerosos estudos sobre cada modelo.

1.1 OBJETIVO

Este trabalho tem o objetivo geral de comparar três abordagens de predição de preços de ações do mercado financeiro, contemplando o cenário mundial de pandemia, a fim de comprovar a ineficiência de modelos altamente recomendados e utilizados em diversos estudos com o mesmo intuito, quando utilizados em cenários incertos.

1.1.1 Geral

Analisar como os três modelos se comportam frente a cenários macro-econômicos incertos.

1.1.2 Específicos

- Definir o processo de obtenção e tratamento de dados.
- Definir os melhores algoritmos para obtenção do resultado desejado.
- Criar modelos computacionais para os algoritmos selecionados.
- Realizar uma análise comparativa dos resultados preditivos.

1.2 ESCOPO DO TRABALHO

O escopo deste trabalho se limita ao estudo comparativo do método estatístico de Simulação de Monte Carlo, modelos de aprendizado de máquina SARIMA e Facebook Prophet, além do modelo de Deep Learning LSTM. Estes modelos computacionais farão a predição de três ações do mercado financeiro brasileiro entre o período de 2015 a 2021, sendo elas: GOLL4 (Gol Linhas Aéreas), PETR4 (Petrobras) e MGLU3 (Magazine Luiza).

1.3 ESTRUTURA DO TRABALHO

Esse trabalho está estruturado de forma que primeiramente temos a introdução, onde são devidamente apresentados o problema e suas características, expondo o contexto por trás do trabalho, assim como seus objetivos gerais e específicos e uma leve apresentação de quais serão os modelos utilizados.

No segundo capítulo, são apresentados todos os assuntos que envolvem esse trabalho e seus modelos, assim como a base teórica das ferramentas utilizadas, como séries temporais, redes neurais, LSTM, entre outros. No próximo capítulo são explicados o desenvolvimento dos modelos, como foram tratados e utilizados os dados, assim como quais foram os resultados obtidos.

Em seguida, o quarto capítulo é dedicado à análise dos resultados em si, realizando uma comparação entre eles e os modelos desenvolvidos. O quinto capítulo apresenta a

conclusão do trabalho e em seguida há o capítulo de trabalhos futuros, apresentando trabalhos que podem ser desenvolvidos a partir da pesquisa realizada. Por fim, há o capítulo onde serão demonstrados trabalhos relacionados.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta teórica e matematicamente cada modelo utilizado, além dos conceitos que esses modelos envolvem.

2.1 SIMULAÇÃO DE MONTE CARLO

A simulação de Monte Carlo consiste em usar dados do passado para criar simulações de vários cenários futuros, calculando a média, variância e desvio padrão. Por meio dessas três métricas conseguimos estimar os preços futuros. Ao fazer as simulações, teremos valores maiores e o objetivo é utilizar os outros menores, enquanto os valores médios são os mais próximos ao valor inicial. O objetivo é utilizar os dados históricos para gerar cenários bons e ruins. Entender o impacto do risco e incerteza ajuda na tomada de decisão, por ser possível responder o que aconteceria em cada situação.

Exemplo, o tomador de decisão poderia traçar um plano de ação caso o pior acontecesse, assim como traçar um plano estratégico para a possibilidade de melhor caso. Para isso, a simulação de Monte Carlo utiliza valores randômicos para gerar os diferentes cenários. O preço de hoje será calculado da seguinte forma:

$$S_t = S_{t-1} \cdot e^r \quad (2.1)$$

Onde r é basicamente, o retorno logarítmico da ação, a taxa de retorno de hoje (preço de hoje/preço de ontem) e S_{t-1} é o preço de ontem.

Para modelarmos r , utilizaremos o movimento Browniano que consiste em duas partes:

a) Drift

Que irá indicar a direção que as taxas de retorno tiveram no passado. Se foram negativas, positivas ou próximas de zero.

$$Drift = \mu - 1/2\sigma^2 \quad (2.2)$$

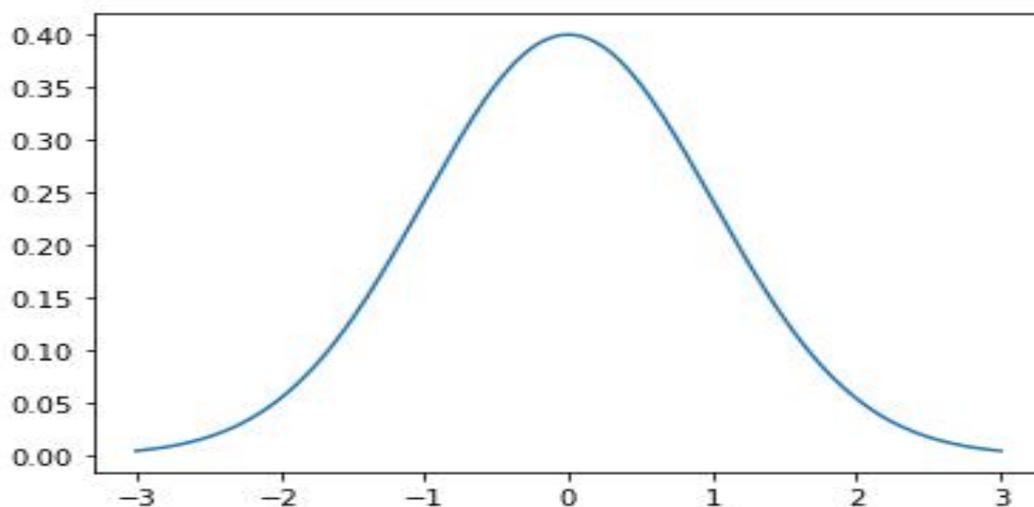
b) Volatilidade

Consideraremos a volatilidade como uma variável aleatória, nós trabalharemos com números randômicos para conseguirmos gerar tanto o melhor cenário quanto o pior cenário.

$$Volatilidade = \sigma Z[Rand(0; 1)] \quad (2.3)$$

Onde Z é considerado uma matriz composta de números randômicos entre 0 e 1. Temos a letra Z , pois estes dados vão precisar estar em uma distribuição normal.

Figura 4 – Exemplo de Distribuição Normal



Fonte: (UNICAMP, 2019)

Com estas informações, conseguimos seguir com o próximo passo que é o cálculo do retorno diário r através do somatório do Drift e Volatilidade.

$$r = (\mu - 1/2\sigma^2) + (\sigma Z[\text{Rand}(0;1)]) \quad (2.4)$$

Por fim, com estes componentes, conseguimos chegar na nossa fórmula de cálculo para o preço de hoje.

$$S_t = S_{t-1}.e^{(\mu-1/2\sigma^2)+(\sigma Z[\text{Rand}(0;1)])} \quad (2.5)$$

2.2 SÉRIES TEMPORAIS

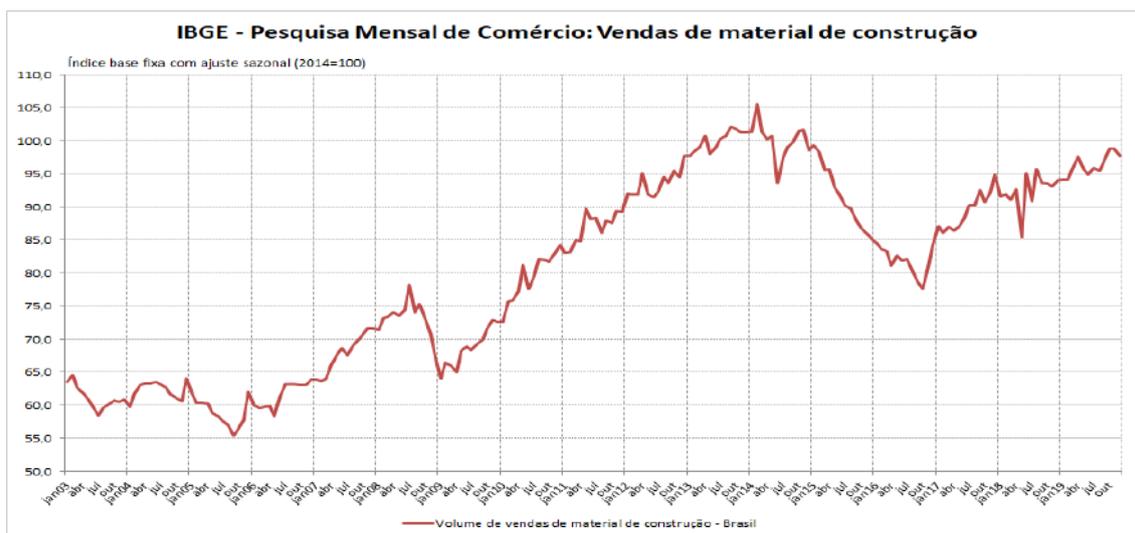
Uma série temporal é o conjunto de pontos (observações) dispostos em ordem cronológica sobre um determinado assunto. Um exemplo de série temporal é o gráfico de vendas de uma empresa em cada mês do ano.

A análise em torno destas séries surge da ideia de que mudanças de um gráfico possuem uma causa e que este fator pode influenciar o gráfico em outro determinado instante. Com isso, este sistema causal tende a atuar criando padrões não aleatórios que podem ser detectados e, conseqüentemente, viabilizar um estudo de causa para a identificação dos fatores que ocasionam a alteração do gráfico. Por fim, com o entendimento dos causadores é possível realizar previsões que irão auxiliar na tomada de decisão de uma empresa ou pessoa.

Analisando a série temporal presente na figura 5 é possível identificar diversos padrões não aleatórios:

1. Existe uma tendência de alta nas vendas, portanto é um negócio que vale a pena o investimento, pois a cada ano tende a vender mais que o anterior
2. Ocorre uma sazonalidade geralmente entre os meses de julho e outubro, evidenciada por uma queda no número de vendas durante o período. Com esta informação, uma empresa produtora poderia produzir menos materiais de construção durante este período para minimizar as perdas.

Figura 5 – Venda de material de construção em função do tempo



Fonte: (FIESP, 2019)

As análises de séries temporais estão longe de serem simples e se tornam cada vez mais complexas em gráficos com padrões complicados de se visualizar, já que análises, como feitas anteriormente, não são suficientes para a tomada de decisão de grandes empresas, que podem ter perdas significativas em caso de erros. Por isso, foram desenvolvidos algoritmos que utilizam aprendizado de máquina para identificar qualquer padrão em uma série temporal.

Portanto, neste trabalho serão utilizados os modelos de séries temporais SARIMA e Facebook Prophet para analisar a variação dos preços das ações de empresas durante a pandemia, verificando a sua assertividade na previsão dos preços e comparando os modelos de forma teórica e prática.

2.2.1 Modelo SARIMA e antecessores

Nesta seção, serão apresentados os conceitos do modelo SARIMA para maior entendimento acerca do tema.

2.2.1.1 Fundamentos teóricos

Este modelo foi desenvolvido em função de outros mais simples chamados de ARMA (Auto-Regressivo e de Média Móvel) de Box e Jenkins (1976) e ARIMA (Auto-Regressivo Integrado de Média Móvel)

O modelo ARMA é composto pelos modelos Auto-Regressivos (AR) e médias móveis (MA).

a) Modelo Auto-Regressivo

Este método é baseado na premissa de que a série observada S_t , possui um determinado ponto observado S_t que pode ser decomposto em uma função das p observações anteriores $S_{t-1}, S_{t-2}, \dots, S_{t-p}$, onde p simboliza o número total de observações feitas antes de S_t .

Este modelo retrata convictamente a ideia da análise de séries temporais observando cada ponto do gráfico e suas respectivas causas, fugindo de acontecimentos aleatórios onde não é possível decompor em um somatório de causas anteriores definidas.

Além desta análise, também são inseridos parâmetros de estruturas autorregressivas σ_i e o ruído branco r_t . Por fim, é possível obter a fórmula final:

$$S_t = \sigma_1 S_{t-1} + \sigma_2 S_{t-2} + \dots + \sigma_p S_{t-p} + r_t \quad (2.6)$$

b) Médias móveis

De acordo com (SOUZA, 2006), nos modelos de médias móveis uma determinada observação S_t pode ser decomposta em uma combinação dos ruídos brancos r_t , ocorridos no período corrente e nos períodos passados.

Se assemelhando bastante com a estrutura de Auto-Regressão, a fórmula final do modelo de médias móveis segue o mesmo conceito, trocando a soma dos valores anteriores por subtração e a substituição do parâmetro de Auto-Regressão pelo da estrutura de médias móveis θ :

$$S_t = r_t - \theta_1 r_{t-1} - \theta_2 r_{t-2} + \dots + \theta_q r_{t-q} \quad (2.7)$$

Com isso, no modelo ARMA, o ponto observado S_t é composto pela combinação dos seus valores passados e do ruído branco ocorrido no período corrente e nos períodos passados, gerando a fórmula final do modelo:

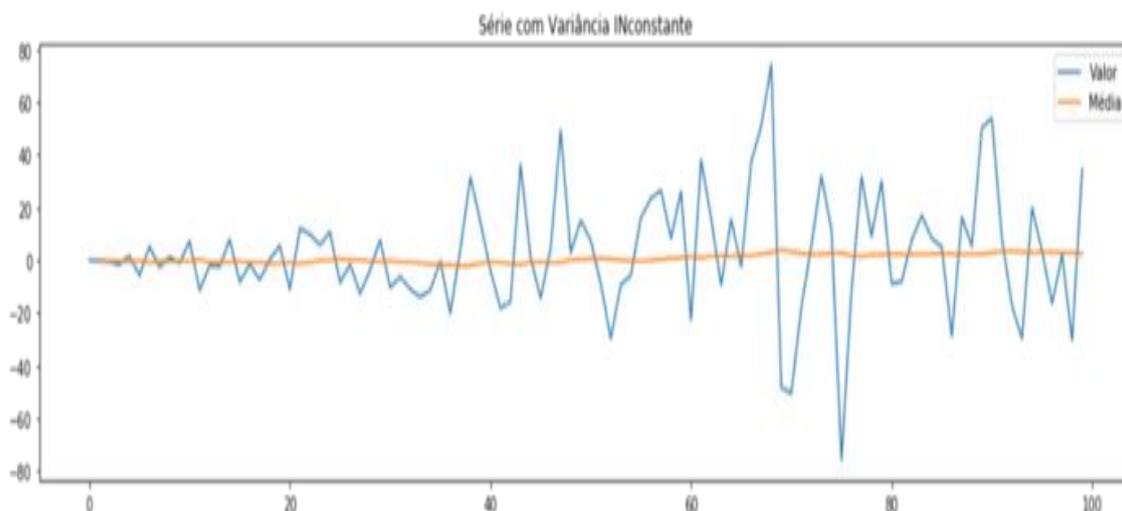
$$S_t = \sigma_1 S_{t-1} + \sigma_2 S_{t-2} + \dots + \sigma_p S_{t-p} + r_t - \theta_1 r_{t-1} - \theta_2 r_{t-2} + \dots + \theta_q r_{t-q} \quad (2.8)$$

No entanto, o modelo ARMA exige que a série temporal seja estacionária para realizar as previsões. Portanto, antes de entrar no próximo modelo, é de suma importância o entendimento sobre séries temporais estacionárias e não estacionárias.

Tipos de Séries temporais

Séries Estacionárias Este tipo de série possui uma média constante durante o período, não ocorrendo nenhuma indicação de crescimento ou decrescimento da média (não existe uma tendência). Estes tipos de séries temporais são mais fáceis de prever, pois os valores vão estar sempre variando ao redor da média (obviamente com alguns ruídos inesperados) A figura 6 Representa o gráfico de uma série temporal estacionária com média constante.

Figura 6 – Venda de material de construção em função do tempo



Fonte: (RABELO, 2019)

Séries não estacionárias Este tipo de série se comporta de forma diferente, a sua média deixa de ser constante e uma tendência é criada. Outros comportamentos não estacionários são ciclos e padrões sazonais (Sendo o foco do modelo ARIMA no comportamento de tendência).

Diante deste problema e sabendo que séries não estacionárias são comuns, foi necessário criar um modelo que conseguisse tratar de séries neste padrão, uma destas formas é transformar a série não estacionária em estacionária realizando o log da mesma e depois subtraindo da média móvel, outra forma é realizando o processo de diferenciação (etapa que será realizada no modelo ARIMA para eliminar tendências). Com isso, surgiu o modelo ARIMA, que, por sua vez, é um dos modelos mais conhecidos e utilizados de séries temporais.

c) Modelo ARIMA

O modelo ARIMA não sazonal é classificado como um modelo “ARIMA (p, d, q)”, em que:

- p é o número de termos autorregressivos (AR);

- d é o número de diferenciações para que a série se torne estacionária (filtro de Integração I);
- q é o número termos de médias móveis(MA).

Uma representação ARIMA(3,2,1), indica um modelo de ordem 3 para o componente AR (Auto-Regressivo), ordem 2 para o componente I (Integração ou diferenciação) e o último 1 para o componente MA (Média Móvel).

Este modelo incorpora os anteriores, sendo a abreviação de autorregressivo, integrado e médias móveis e segue a metodologia Box-Jenkins para modelagem de seu processo, seguindo 4 etapas importantes:

Identificação A etapa consiste em identificar os valores de (p, d, q) segundo a série temporal em análise. Para isso, a primeira etapa é identificar se a série é estacionária ou não. Se ela for estacionária, o filtro de integração terá ordem zero e os valores passariam a ser $(p,0,q)$, fazendo com que o modelo tenha o mesmo comportamento do ARMA. Porém, caso contrário, serão realizadas diferenciações para torná-la estacionária, em caso de 1 diferenciação torná-la estacionária o valor de “ d ” será 1 (O nome deste processo é Diferenciação).

No processo de diferenciação é calculada a diferença entre um valor em um dado momento e o valor anterior, gerando um novo gráfico que terá uma média mais constante.

Para melhor entendimento, segue o exemplo:

A figura 7 representa um gráfico não estacionário, a figura 8 representa os valores antes e depois do processo de diferenciação e a figura 9 é o gráfico final após a diferenciação. Neste caso, podemos observar que apenas uma diferenciação foi necessária para transformar o gráfico em estático.

Após a identificação do parâmetro “ d ”, são utilizados os algoritmos de autorregressão e médias móveis para descobrir o valor das outras variáveis.

Estimação Após encontrar os valores adequados de p, d e q , chega o processo de estimação dos parâmetros dos termos autorregressivos e de médias móveis incluídos no modelo. Esta etapa é bastante complexa e existem diferentes formas de ser feita, e são geralmente utilizados softwares específicos para este procedimento.

Validação Nesta etapa, o modelo encontrado é validado para garantir que será funcional para a série estudada. Uma técnica muito utilizada para esta etapa é a análise de resíduos, pois, a partir dos mesmos, teremos ideia se as suposições feitas sobre o modelo são ou não válidas. Caso seja comprovado que o modelo proposto é adequado, o algoritmo segue para a próxima e última etapa, se não, deve retornar para as anteriores e recalculer os parâmetros.

Previsão Por fim, com o modelo adequado escolhido, é possível começar a realizar previsões acerca da série temporal inicial.

Figura 7 – Série não estacionária

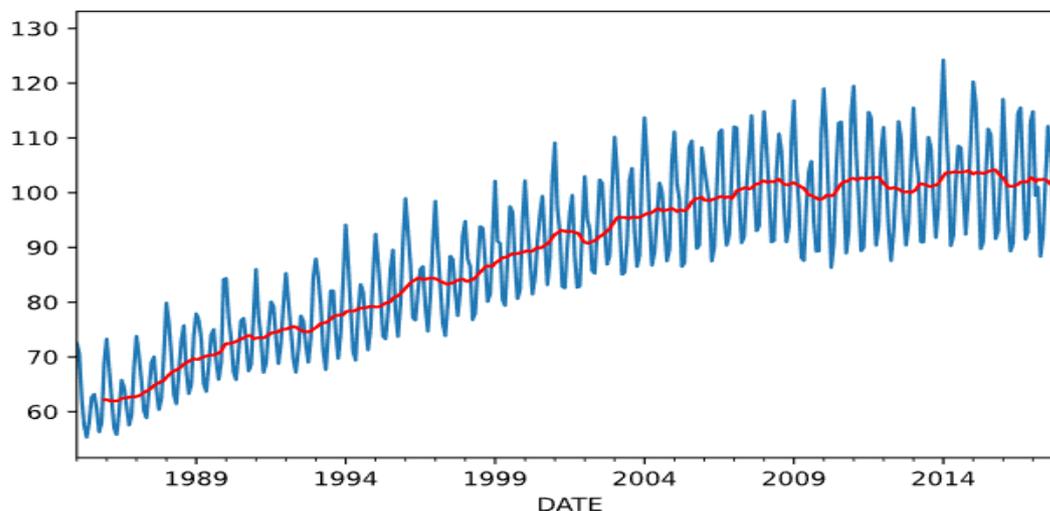


Figura 8 – Valores antes e depois da diferenciação

```

Valores Antes da diferenciação
<bound method NDFrame.head of
DATE                                     Value
1985-01-01    72.5052
1985-02-01    70.6720
1985-03-01    62.4502
1985-04-01    57.4714
1985-05-01    55.3151
...
2017-09-01    98.6154
2017-10-01    93.6137
2017-11-01    97.3359
2017-12-01   114.7212
2018-01-01   129.4048

[397 rows x 1 columns]>
Valores Depois da diferenciação
<bound method NDFrame.head of
DATE                                     Value
1985-01-01      NaN
1985-02-01    -1.8332
1985-03-01    -8.2218
1985-04-01    -4.9788
1985-05-01    -2.1563
...
2017-09-01   -10.3158
2017-10-01    -5.0017
2017-11-01     3.7222
2017-12-01    17.3853
2018-01-01    14.6836

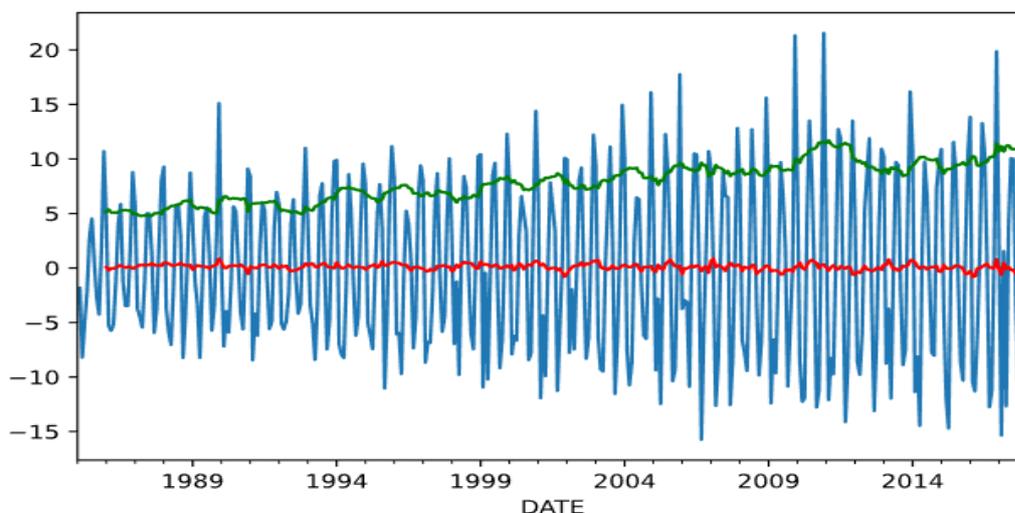
```

Fonte: Imagem gerada via código

d) **Modelo SARIMA** Este modelo surge de uma diferença entre séries estacionarias e não estacionarias que não é coberta pelo algoritmo ARIMA, as sazonalidades.

O modelo ARIMA sazonal ou SARIMA, segue o seguinte modo de apresentação $ARIMA(p, d, q)(P, D, Q)$, onde os parâmetros (P, D, Q) são os equivalentes sazonais de (p, d, q) . De acordo com (FAVA, 2000) os instrumentos utilizados na

Figura 9 – Gráfico após diferenciação



identificação de modelos sazonais continuam sendo a FAC (Função Autocorrelação) e FACP (Função Autocorrelação Parcial).

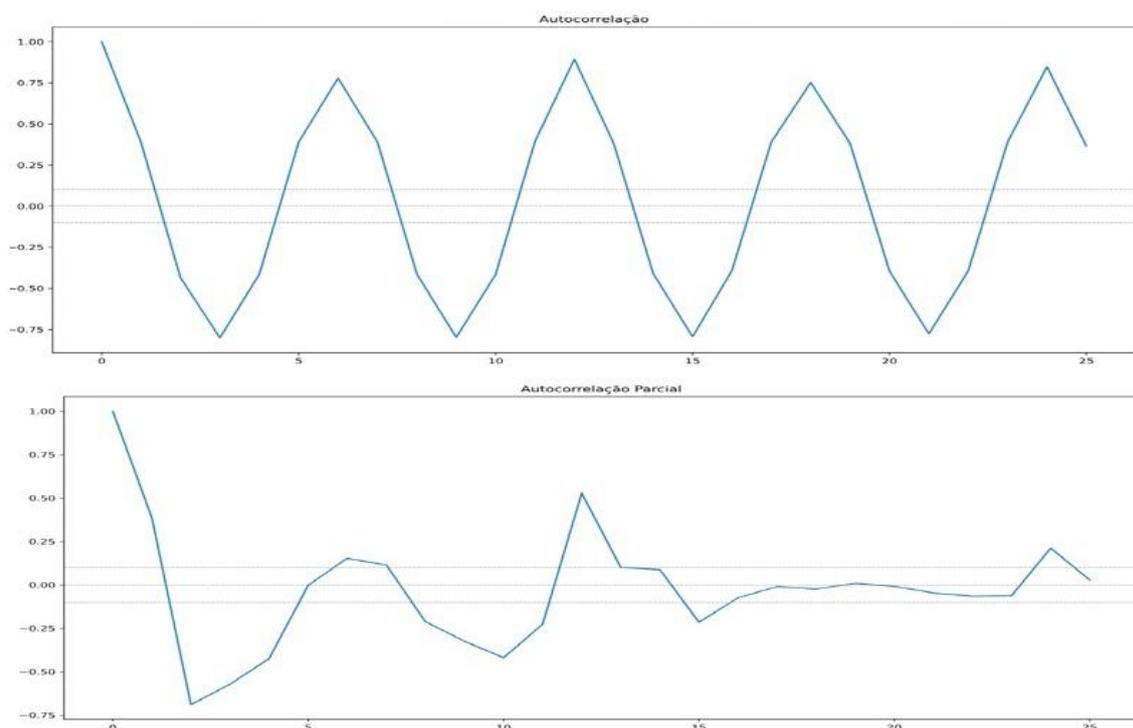
O fator sazonalidade é referente a uma repetição de padrão na série temporal em um período específico. Por exemplo, uma loja que sempre possui uma alta nas suas vendas no período de janeiro até junho. Portanto, o algoritmo de previsão precisa levar isso em conta, pois mesmo que exista uma tendência de queda nas vendas, se estiver entrando em janeiro, deve prever que mesmo em queda em uma visão macro, existirá um aumento momentâneo no número de vendas da loja.

O primeiro passo do algoritmo é determinar os parâmetros d e D , que pode ser encontrado com a análise da FAC da série original e de suas diferenças, consecutivas, sazonais ou ambas. Após as diferenciações d e D , é necessário determinar os outros parâmetros p , P , q e Q . Estes demais parâmetros podem ser encontrados através da análise de ACF e PACF.

Esta análise é uma técnica gráfica que ajuda a identificar as ordens apropriadas dos componentes AR e MA em um modelo SARIMA. Consideremos o seguinte exemplo: A Figura 9 representa uma série temporal após a diferenciação, pois, como foi explicado anteriormente nos outros modelos, para identificar os valores de p e q a série precisa ser estacionária. Com a Utilização da linguagem Python, foi possível gerar a Função de Autocorrelação e a Função de Autocorrelação Parcial representadas na figura 10.

Através da Observação dos gráficos de autocorrelação é possível obter os valores de q e p . Antes de tudo, é importante salientar o significado das linhas horizontais nos gráficos, tais linhas representam a significância dos valores, portanto os valores importantes são os que se encontram acima ou abaixo da linha. Com isso, é possível identificar os seguintes fatores:

Figura 10 – gráficos de autocorrelação e autocorrelação parcial



q → valor onde o gráfico de autocorrelação cruza o limite superior do intervalo pela primeira vez, neste caso seria o 2 do eixo X.

p → valor onde o gráfico de autocorrelação parcial cruza o limite superior do intervalo pela primeira vez, neste caso seria também o 2 do eixo X.

Portanto, foram obtidos os valores-base $(p, d, q) = (2, 1, 2)$ possibilitando a realização de previsões no modelo ARIMA cuja análise de seus resíduos gerados (diferença entre os valores observados e os valores previstos pelo modelo) possibilita uma nova análise da FAC e dos gráficos de autocorrelação em cima destes resíduos para encontrar as componentes sazonais (P, D, Q) .

Depois desta fase de identificação o algoritmo segue para as próximas etapas explicadas anteriormente.

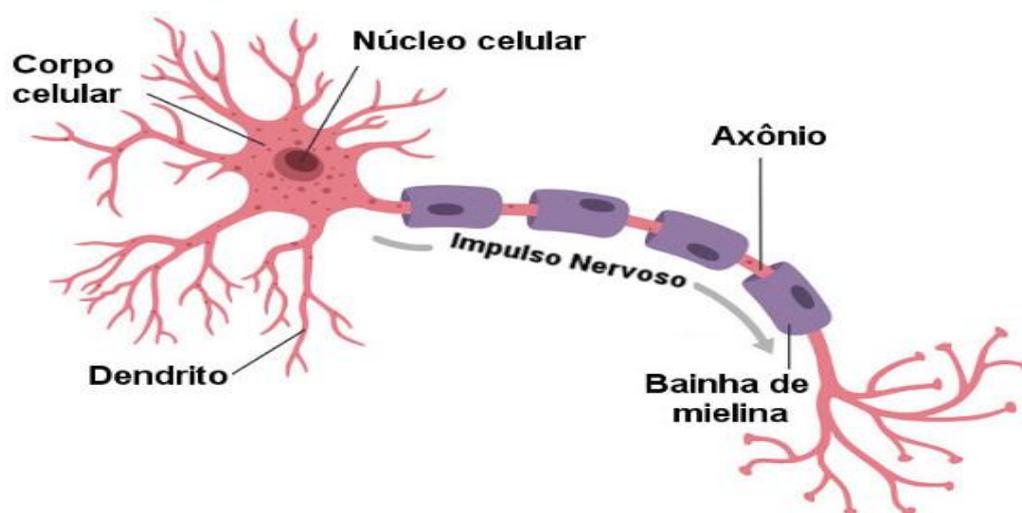
2.3 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA), ou simplesmente Redes Neurais, são estruturas computacionais criadas para replicar o funcionamento de neurônios biológicos. A RNA visa identificar certas tendências ou padrões em um conjunto de dados e a partir disso, efetuar uma generalização. A maior vantagem em utilizar essa estrutura está na sua capacidade de aprender esses padrões subjacentes presentes nos dados, sendo onde a maioria dos métodos convencionais costuma falhar (ZHANG; PATUWO; HU, 1998). Na

visão da biologia, neurônios são as células responsáveis pela transmissão de impulsos nervosos provenientes de estímulos vindo do ambiente externo ou do próprio organismo.

Como pode ser visto na Figura 11, o neurônio é dividido basicamente em três partes: dendritos, axônios e corpo celular. Os dendritos são terminações que captam informações externas e garantem a recepção de estímulos, dessa forma, levando impulsos nervosos até o corpo celular, que por sua vez, é onde está presente o núcleo e onde serão processadas as informações que chegam dos dendritos. No axônio, é onde os impulsos são conduzidos e chegam a outros neurônios. Há também uma estrutura denominada bainha de mielina no axônio, que confere um isolamento elétrico a essa região, não permitindo a perda de informações.

Figura 11 – Estrutura de um neurônio



Fonte: (SARDINHA, 2019)

Os neurônios possuem a capacidade de receber diversas informações ao mesmo tempo, e processá-las, emitindo uma única resposta, e é esse comportamento que as redes neurais pretendem replicar. Para isso, são utilizados neurônios artificiais, que nada mais são que unidades de processamento de informações. Esses neurônios artificiais se conectam por meio de um sistema de pesos atribuídos de acordo com seus aprendizados nos dados (MOGHADDAM; MOGHADDAM; ESFANDYARI, 2016) e utilizam uma função de ativação para retornar uma saída (HAYKIN, 2001).

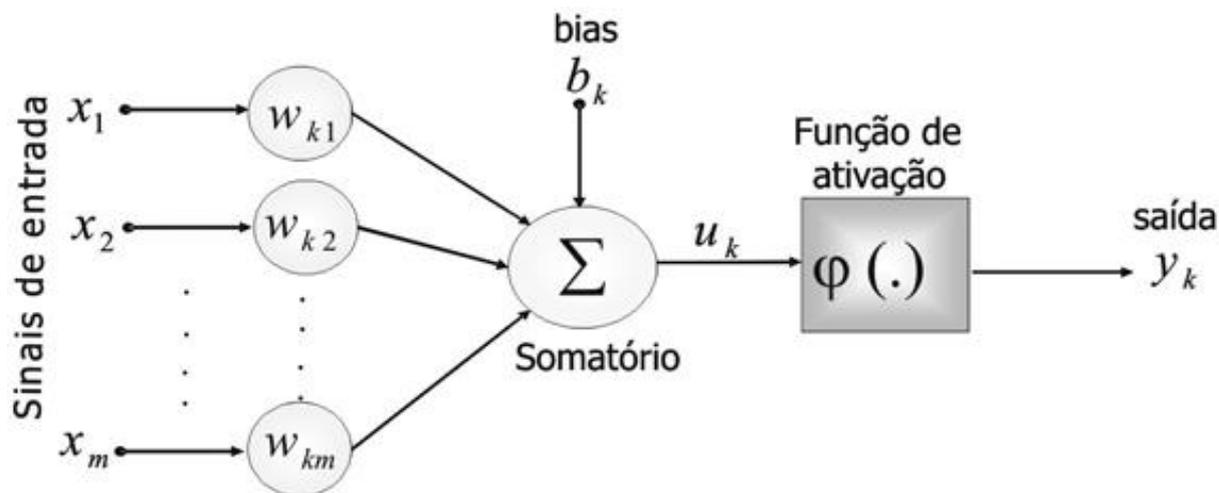
A conexão entre os neurônios artificiais recebe o nome de sinapse, que nada mais é por onde as informações serão passadas de um neurônio ao outro. Essa informação que chega como entrada a um neurônio é multiplicada por um peso definido por quem construiu aquela rede. Esse peso é o que ajusta a importância dos diferentes resultados computacionais que o neurônio obterá.

Na Figura 12, temos a representação de um neurônio artificial. Ele irá receber m entradas (x_m), que serão conectadas ao neurônio por sinapses com pesos atribuídos (w_{km}).

Essas entradas são somadas pelo neurônio utilizando a equação:

$$u_k = \sum_{i=1}^m x_i w_{ki} + b_k \quad (2.9)$$

Figura 12 – Modelo de um neurônio artificial



Fonte: Adaptado de (HAYKIN, 2001)

Também há um valor de viés (b_k), sendo a distância dos valores previstos para os valores reais, ou seja, o erro obtido no modelo. Esse somatório retorna um valor u_k , que é, então, aplicado a uma função de ativação para produzir a saída final.

Existem diversos tipos de RNA, além de que a todo momento são desenvolvidas novas variações dos modelos já existentes. Alguns desses tipos podem ser observados na Figura 13. É claro que essas redes não se diferenciam apenas pela sua estrutura. Alguns outros fatores também são considerados, como o método de treinamento delas. As redes mais importantes para o entendimento deste trabalho serão apresentadas posteriormente.

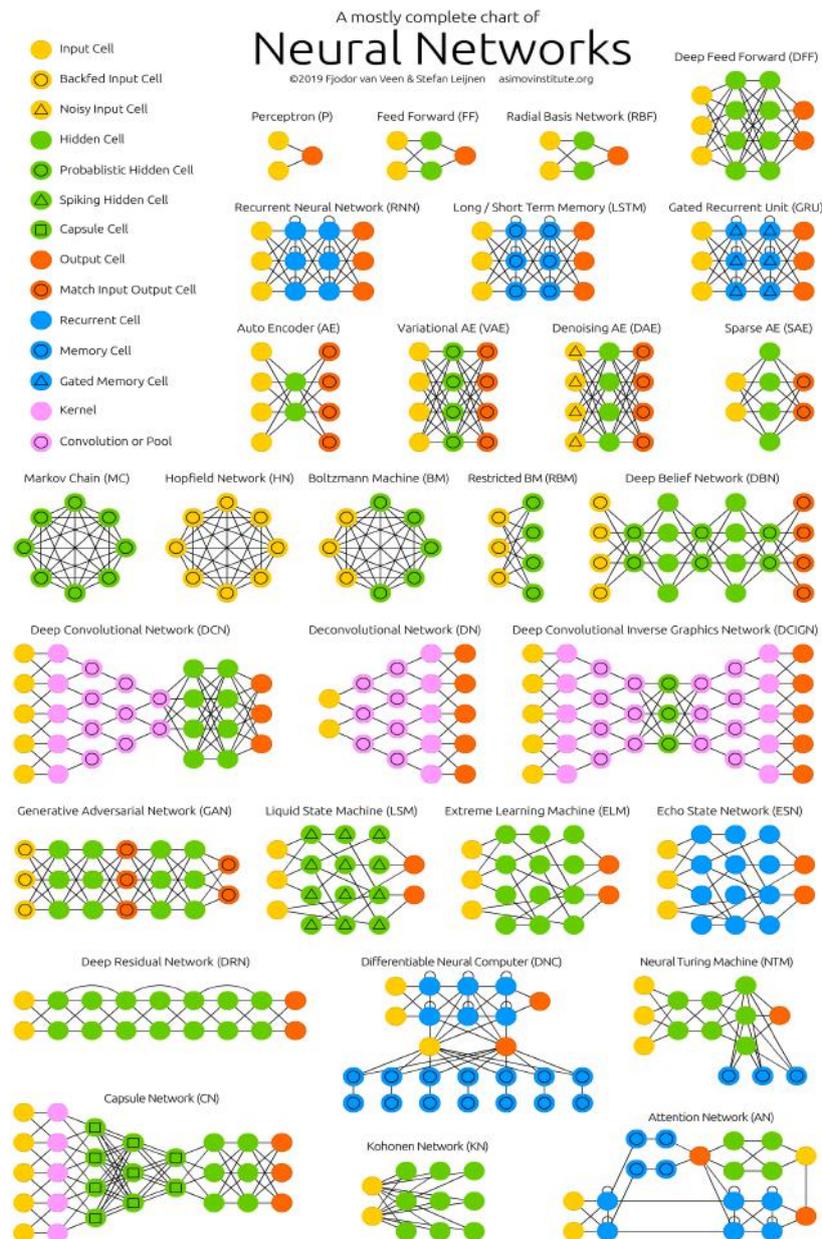
2.3.1 Função de Ativação

As funções de ativação, como visto anteriormente, estão em cada neurônio de uma RNA e apesar de existir a possibilidade de se utilizar uma função linear, o ideal é o uso de funções não-lineares, pois além de dados financeiros também serem não-lineares, a função de ativação também induz uma não linearidade na saída do neurônio e tornam essa estrutura capaz de aprender mais do que relações lineares entre variáveis dependentes e independentes. Nesta seção, serão apresentadas as duas funções mais utilizadas.

a) Função Sigmoid

A equação da função sigmoide é a seguinte:

Figura 13 – Tipos de Redes Neurais



Fonte: (VEEN, 2016b)

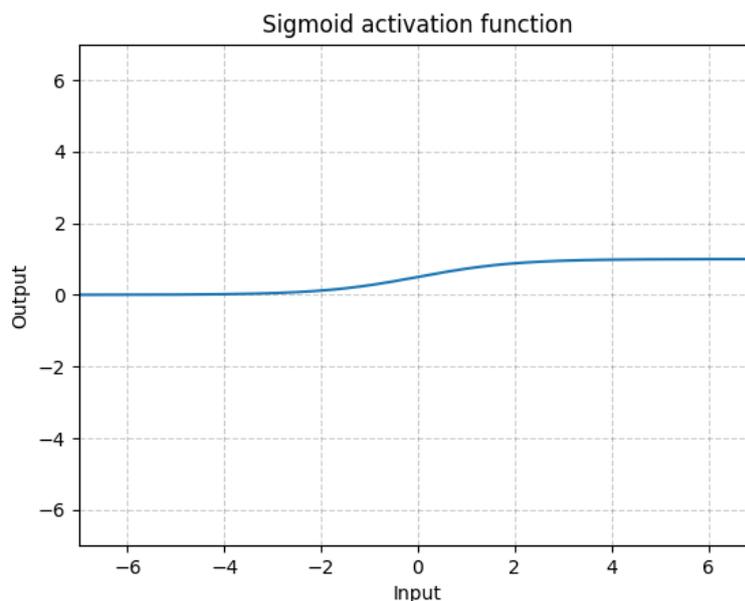
$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.10)$$

A função sigmoide é a primeira das funções classificadas como sigmoidal, devido ao seu formato, que lembra a letra S. Ela também foi a primeira função a ser utilizada nessa aplicação de redes neurais (WLODZISLAW; JANKOWSKI, 1999). Como o intervalo dessa função está entre 0 e 1, ela é normalmente utilizada em classificações binárias, onde um resultado maior que 0,5 é previsto como 1 e abaixo

desse valor, é previsto como 0.

O gráfico dessa função pode ser visto na Figura 14.

Figura 14 – Gráfico da Função Sigmoide



Fonte: PyTorch

b) Função Tangente Hiperbólica

A equação da função tangente hiperbólica é a seguinte:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.11)$$

Essa é outra função sigmoide, utilizada em casos parecidos aos da função sigmoide. Porém, agora o intervalo de valores dessa função está entre -1 e 1, facilitando na centralização dos dados, já que as médias serão 0 ou valores muito próximos disso. A função tanh é uma versão matematicamente alterada da função sigmoide.

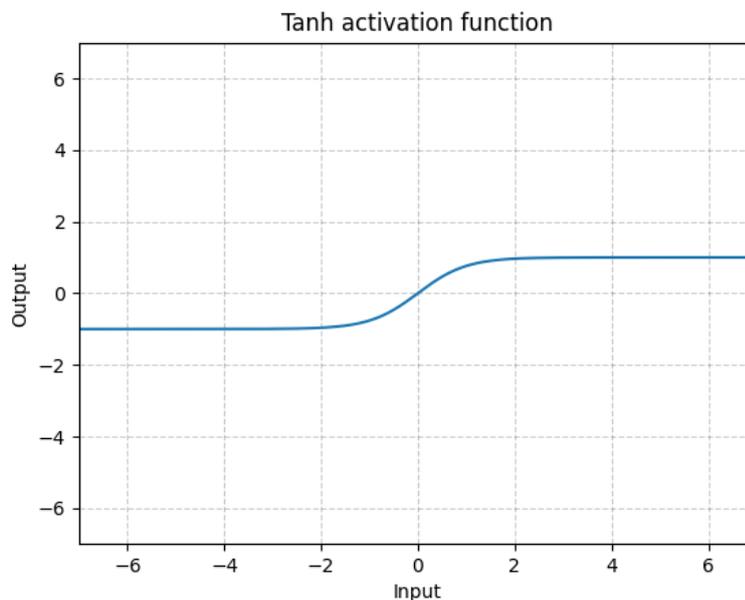
Na Figura 15, pode ser visualizado o gráfico dessa função.

2.3.2 Arquitetura

A arquitetura de uma RNA diz respeito à maneira que neurônios são agrupados e conectados. Essas configurações podem ser modificadas e combinadas para diferentes aplicações.

Uma rede neural pode possuir apenas uma ou diversas camadas com neurônios, onde cada um desses neurônios recebe a entrada de outro neurônio de uma camada anterior e a sua saída chega a um neurônio da camada posterior, porém o caminho inverso também

Figura 15 – Gráfico da Função Tangente Hiperbólica



Fonte: PyTorch

pode ser realizado. Como poder ser visto na Figura 16, uma RNA é composta pela camada de entrada, onde os dados chegam. as camadas ocultas, também chamadas de camadas intermediárias ou escondidas e por fim, a camada de saída.

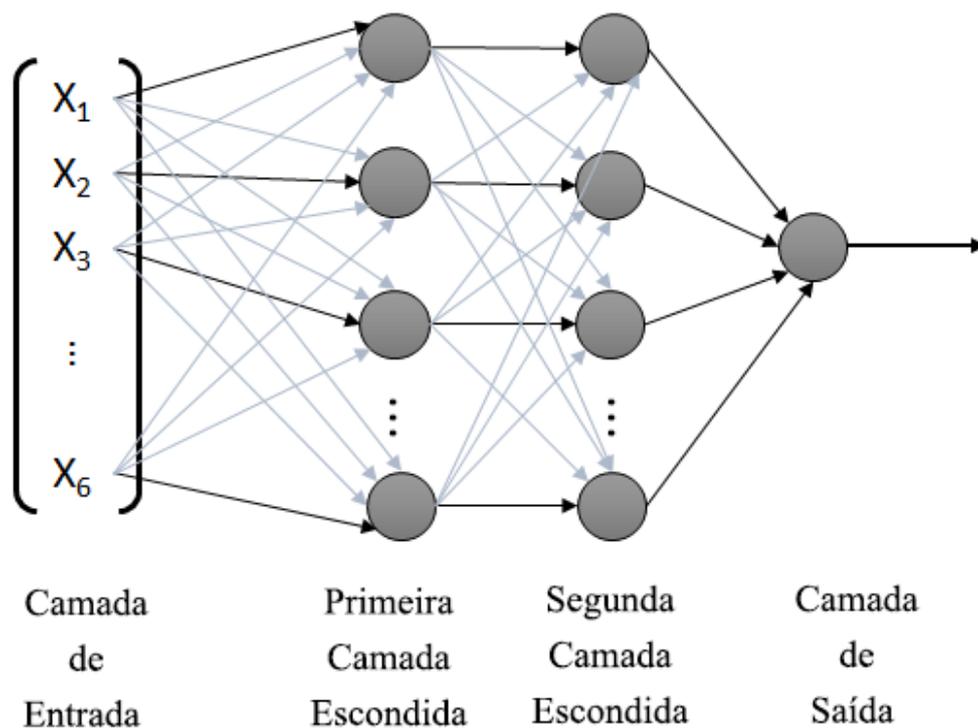
Para redes com múltiplas camadas, existem 3 padrões de conectividade entre seus neurônios (HAYKIN, 2001). Há o caso completo, onde cada neurônio está conectado a todos os anteriores e posteriores; o local, onde os neurônios que um determinado neurônio irá se conectar são definidos em uma região da camada; e o parcial, onde o neurônio também se conecta somente a alguns outros, mas isso não necessariamente se aplica apenas a neurônios da mesma região.

Conforme a quantidade de neurônios nas camadas de entrada e saída, a arquitetura também pode receber diferentes nomenclaturas. Se há apenas uma entrada e uma saída, é dito que é uma arquitetura um para um, enquanto para múltiplas entradas e múltiplas saídas, tem-se a arquitetura muitos para muitos. O raciocínio segue para as permutações dessas ocasiões (um para muitos e muitos para um) (KARPATHY, 2015).

2.3.3 Processo de Aprendizado

Uma rede neural irá utilizar os dados passados para gerar um modelo, tornando a fase de aprendizado extremamente importante. Segundo (GIRIOLI; RIBEIRO, 2009), de 50% a 90% do total de dados devem ser separados para treinamento da rede, além de que esses dados devem ser escolhidos aleatoriamente, fazendo com que a rede de fato aprenda as regras e não decore exemplos. Os dados que sobram só serão apresentados à rede neural

Figura 16 – Exemplo de arquitetura de uma RNA com duas camadas ocultas



Fonte: (VIDAL et al., 2015)

na fase de testes, com isso a rede irá ter um maior poder de dedução da relação entre os dados.

Deve-se escolher um algoritmo de aprendizado, sendo os mais tradicionais o algoritmo supervisionado e o não supervisionado. Outra escolha a ser efetuada é como as informações serão transmitidas na RNA, um exemplo é a presença ou não de retroalimentação na rede, assunto esse que será discutido posteriormente.

Antes do treinamento em si, há uma fase de configuração da rede neural, feita com base em observações e testes realizados pelo próprio usuário até atingir a melhor configuração para o problema tratado. Com isso, ocorre então o treinamento, onde os pesos são escolhidos. O tempo do treinamento poderá variar segundo a escolha desses pesos e ele deve ser interrompido assim que a rede demonstrar uma boa capacidade de generalização e o erro retornado for suficientemente pequeno.

Por último, ocorre a fase de testes, onde o conjunto de dados anteriormente separado para testes serão finalmente utilizados para visualizar a atuação da rede neural. O resultado aqui obtido é o que irá mostrar como a rede irá, de fato, trabalhar em uma situação real.

2.3.3.1 Aprendizado Supervisionado

No tipo de aprendizado denominado supervisionado, há uma saída desejada comparada com a saída atual da rede, assim ela tem conhecimento do quão apurado está seu resultado e pode ajustar o valor dos pesos, diminuindo o erro obtido no treinamento e obtendo um resultado mais preciso. Esse método recebe o nome de retropropagação ou backpropagation.

Outro fator a ser considerado é quando acontecerá o processo de mudança nos pesos. O método de treinamento mais utilizado, é o treinamento por lote, ou ciclo, que requer o processamento de toda a entrada para calcular essa mudança nos pesos. A cada fim de ciclo que o algoritmo realiza para reajustar os pesos e aumentar a precisão da RNA, ocorre o fim do que é chamado de época.

O valor utilizado para ajustar o peso do neurônio, seja somando ou subtraindo dele, é denominado taxa de aprendizado e esse valor traz algumas interpretações sobre o aprendizado da rede. Se a taxa de aprendizado for muito alta, ocorrerá uma mudança de pesos maior e o algoritmo acabará ultrapassando o melhor peso de cada neurônio. Já para uma taxa muito pequena, o algoritmo acabará demorando demais até finalmente encontrar um bom valor para os pesos (LA, 2019).

2.3.3.2 Aprendizado Não Supervisionado

No aprendizado não supervisionado, não há mais uma saída desejada, o que faz com que a rede realize uma compressão nos dados para agrupar valores que constituem um padrão. Um padrão é reconhecido mediante uma seção de treinamento onde são apresentados um conjunto de padrões de entrada e suas categorias repetitivamente. Após isso, é apresentado um padrão nunca visto à rede, mas ainda pertencente a população dos padrões que já foram vistos. Com a informação que a rede obteve no aprendizado, a rede consegue identificar a categoria correta do padrão (VASCO, 2020).

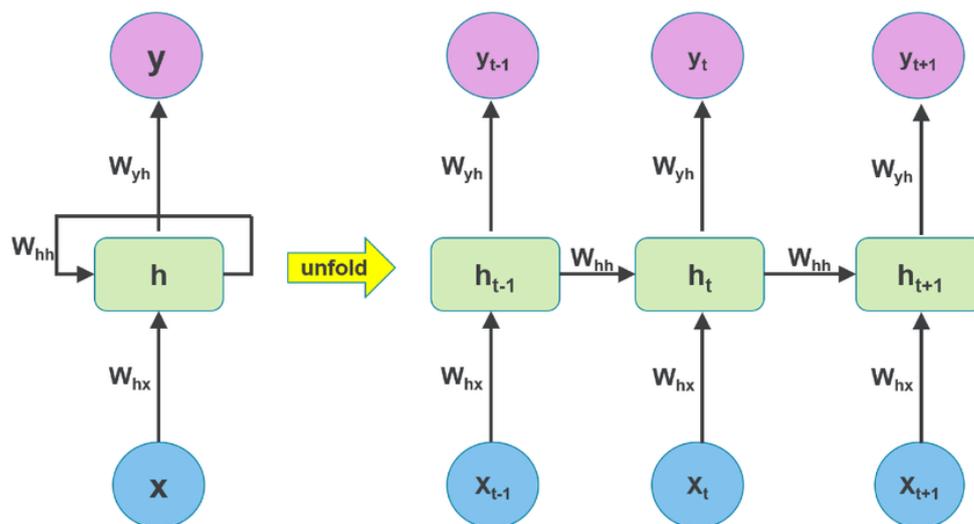
2.3.4 Rede Neural Recorrente

Para o entendimento da rede neural em destaque nesse trabalho, é preciso antes entender o que é uma rede neural recorrente. Ao apresentar o que é uma RNA, os conceitos e exemplos debatidos seguiram o modelo de uma rede neural conhecida como feedforward, a qual é uma das classes mais básicas de redes neurais, ficando na frente apenas da rede conhecida como Perceptron. Nesse tipo de rede, os dados são passados pelas camadas de forma contínua, sem a presença de loops. Basicamente, a informação entra na rede, é processada e transformada em saída, seguindo sempre para frente e nunca enviada de volta, como é o sentido de propagação visto na Figura 16.

As redes neurais recorrentes apresentam loops em sua estrutura, ou seja, a entrada de um neurônio é realimentada com sua própria saída, permitindo o armazenamento de

valores ao longo do tempo. Os dados são armazenados nas camadas intermediárias, que se conectam à camada intermediária do neurônio seguinte até chegar a camada de saída. A estrutura dessa rede neural pode ser vista, de forma simples, na Figura 17, onde x é a entrada, y a saída, h é a camada oculta e W representa os pesos da rede. Por tratar de problemas não lineares, as funções de ativação mais utilizadas são Sigmoid e Tanh.

Figura 17 – Arquitetura simples de uma RNN com um único neurônio e com três



Fonte: (VENUGOPAL; T., 2009)

Se olharmos para a rede com mais neurônios (parte direita da figura) e considerarmos f uma função de ativação, vemos que a equação da camada intermediária é dada por:

$$h_t = f(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \quad (2.12)$$

Enquanto para a saída, temos que:

$$y_t = f(W_{yh}h_t + b_h) \quad (2.13)$$

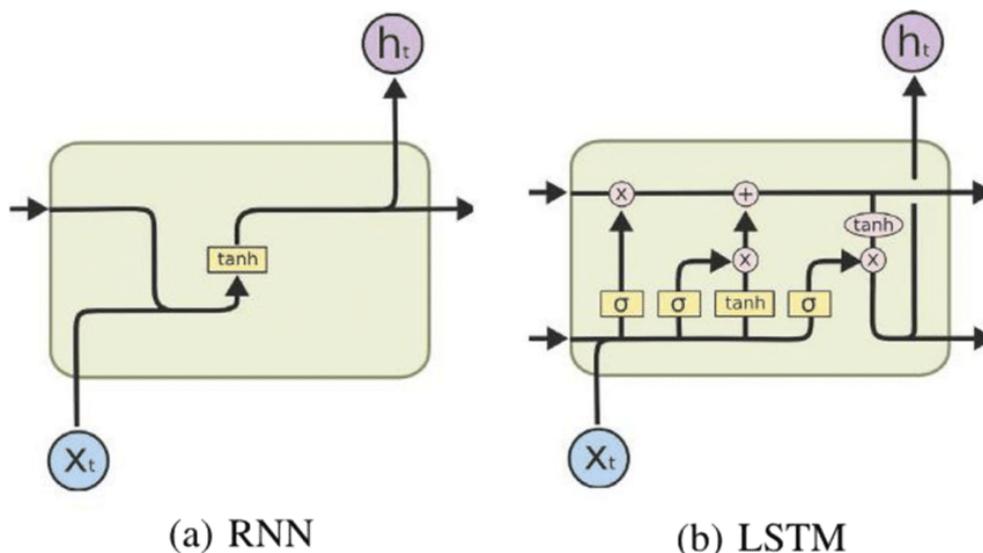
Esse comportamento torna a rede capaz de trabalhar com certas aplicações, como ao tratar dados sequenciais e também replica uma memória artificial, se assemelhando ainda mais com a forma de processamento de informação das redes neurais humanas. Essa adição de memória faz com que cada dado inserido afete ele mesmo e os dados que serão inseridos após ele, já que ocorrerá uma comparação entre eles. É essa característica que torna esse tipo de rede essencial para a realização de uma predição, já que dados do passado irão afetar diretamente os dados do futuro e por conta disso, a ordem com que esses dados chegam também é importante. Outros exemplos de utilização dessa rede são problemas temporais, processamento de linguagem natural, reconhecimento de fala, tradução de linguagem.

Um problema das redes neurais recorrentes é que as informações de curto prazo recebem maior importância para o cálculo da camada atual, deixando de lado camadas anteriores, além de que o gradiente da função de ativação decai exponencialmente com o passar do tempo, fazendo com que o modelo perca cada vez mais sua capacidade de aprender, levando o nome de problema da dissipação do gradiente (HAYKIN, 2009). Para trabalhar com dados de um passado mais antigo, existe a rede neural conhecida como Long Short-Term Memory.

2.3.5 Rede Long Short-Term Memory

A rede Long Short-Term Memory(LSTM) é um tipo de rede recorrente que consegue aprender dependências de longo prazo. Sua principal característica é que ela consegue aprender o que armazenar no estado de longo prazo, o que jogar fora e o que ler a partir dele, a fim de evitar os problemas mencionados na seção anterior. Esse tipo de rede é recomendada para classificação e previsão de dados de séries temporais. A arquitetura da camada oculta de uma célula LSTM pode ser vista na Figura 18 ao lado de uma célula RNN.

Figura 18 – Comparação entre células RNN e LSTM

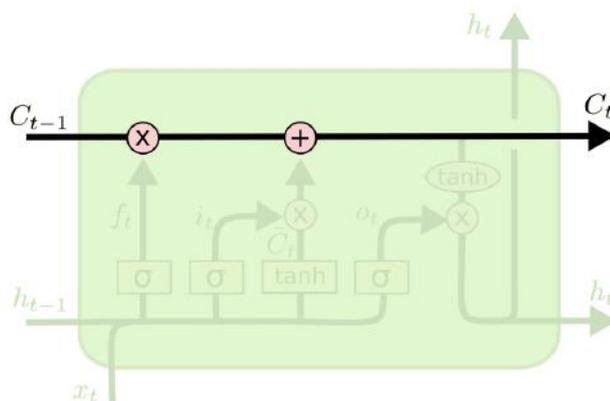


Fonte: (MAMANDIPOOR et al., 2020)

Como pode ser visto, a célula da rede LSTM possui uma estrutura mais complexa. Ela é dividida em algumas partes. Uma delas é o chamado estado da célula, representado na Figura 19. Por ali irá passar a informação, que sofrerá algumas operações pontuais. Essa informação é adicionada ou removida do estado da célula por estruturas internas denominadas portões, que agem como uma memória, definindo o que a rede irá guardar para si e o que irá esquecer ao longo do tempo. Esses portões possuem uma camada

sigmoide que retorna valores entre 0 e 1. Esses valores definem o quanto da informação continuará sendo utilizado.

Figura 19 – Estado da célula — LSTM

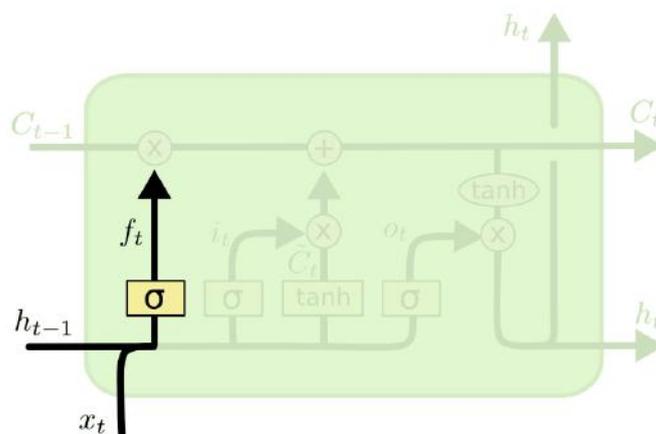


Fonte: (OLAH, 2015)

Dentro de cada célula, há 3 portões de controle. A primeira etapa do fluxo é passar pelo primeiro portão, conhecido como Forget Gate. É este portão que define o que é esquecido pela rede através da camada sigmoide. Para valores próximos a 0, mais dados são esquecidos e para valores próximos de 1, mais dados são mantidos. Ele pode ser observado na Figura 20 e sua equação é a seguinte:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.14)$$

Figura 20 – Forget Gate — LSTM



Fonte: (OLAH, 2015)

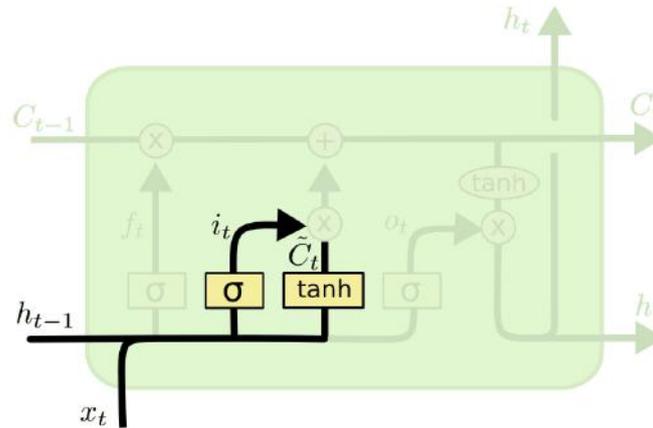
Após isso, a célula realiza a verificação de qual dado será armazenado no estado através do Input Gate. Esse portão também possui uma função sigmoide para decidir o que é importante ser armazenado no estado da célula, além de uma função tangente hiperbólica

para definir o que será acrescentado ou subtraído. As equações desse portão são vistas abaixo, assim como a Figura 21, com ele representado.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.15)$$

$$\vec{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.16)$$

Figura 21 – Input Gate — LSTM



Fonte: (OLAH, 2015)

Agora, o próximo passo é atualizar o estado da célula considerando os dois passos anteriores. Aqui, realiza-se a multiplicação do estado antigo pelo coeficiente do Forget Gate, definindo o quanto será mantido dessa informação, e a adição dos novos valores vindos do Input Gate, definindo quais informações serão atualizadas. Essa etapa pode ser vista na Figura 22, assim como sua equação a seguir:

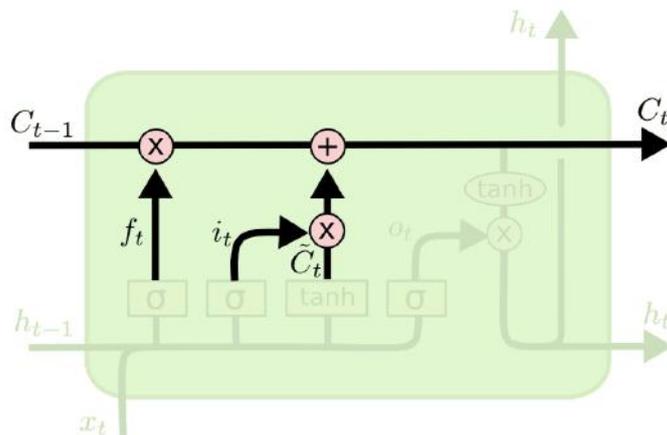
$$C_t = f_t \times C_{t-1} + i_t \times \vec{C}_t \quad (2.17)$$

Por último, há o Output Gate, que possui um filtro para controlar o quanto da informação do estado atual é utilizado para determinar a saída da célula. Aqui, também há o produto de uma função sigmoide com uma tanh, porém agora, os valores que passam pela camada da tangente hiperbólica vêm do estado da célula atualizado. Esse comportamento é visualizado nas equações abaixo e na Figura 23.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.18)$$

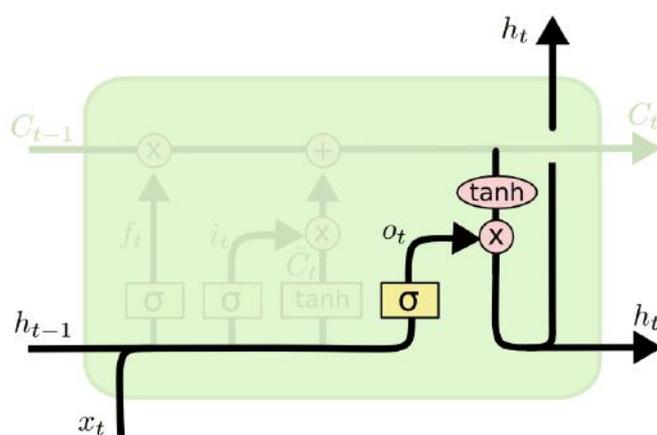
$$h_t = o_t \times \tanh(C_t) \quad (2.19)$$

Figura 22 – Atualização do estado da célula



Fonte: (OLAH, 2015)

Figura 23 – Output Gate — LSTM



Fonte: (OLAH, 2015)

3 TRABALHOS RELACIONADOS

A predição de séries temporais financeiras é um campo de grande interesse na área de inteligência artificial aplicada ao mercado financeiro, tanto no meio acadêmico quanto no setor corporativo. Um trabalho que também é voltado para essa área é o realizado por (VASCO, 2020). Nele, também são comparados três modelos ao prever valores de ações, porém, no caso, são três modelos de redes neurais recorrentes, sendo uma dessas redes a LSTM, também utilizada neste trabalho. As outras redes utilizadas foram o modelo de redes neurais recorrentes simples (VRNN, do inglês Vanilla Recurrent Neural Network) e o Gated Recurrent Unit (GRU).

O estudo utilizou dados de três ações diferentes da bolsa de valores brasileira, mais precisamente as ações do Grupo Fleury (FLRY3), da Petrobras (PETR4) e da Vale (VALE3). Para cada uma das ações, foram criados modelos baseados nas respectivas arquiteturas de redes neurais recorrentes para realizar a predição dos valores dessas ações. O objetivo principal foi comparar os resultados obtidos por cada modelo para avaliar qual predição se aproximou mais dos valores reais. Para a avaliação do desempenho dos modelos, foram utilizadas medidas de erro, como o erro absoluto médio e o erro quadrático médio. Os resultados indicaram que os modelos GRU e LSTM apresentaram resultados semelhantes, com uma ligeira vantagem para o desempenho do GRU em relação ao LSTM. Por outro lado, o VRNN também pôde detectar padrões nos dados, mas sua precisão não foi tão alta em comparação com os outros modelos.

No ramo das séries temporais existem, também, outros artigos abordando o modelo de predições SARIMA. No entanto, todos os trabalhos encontrados testavam o modelo em situações mais controladas e terminavam com grandes desempenhos nas predições, conforme realizado no artigo (ESPINOSA, 2010). Com isso, foi escolhido o modelo na esperança de ter ótimos resultados também em predições de mercado, mas o resultado foi longe do esperado. Até foi considerado trocar o modelo por outro que iria ter melhores resultados, porém, no início deste trabalho, foi levantado a possibilidade de ter resultados que não são esperados e que isso não significa um problema para um artigo ou trabalho de conclusão, pois resultados negativos não deixam de ser resultados e eles sempre irão vir com informações importantes que podem acrescentar e muito no modelo.

4 DESENVOLVIMENTO DA PROPOSTA

Este capítulo apresenta como os modelos foram utilizados para a obtenção dos resultados, além dos cenários escolhidos para cada um e os parâmetros utilizados.

4.1 DADOS

Nesta seção, serão apresentados os dados utilizados no trabalho, trazendo informações como de onde foram retirados, quantidade de dados e a finalidade de escolher esses três conjuntos de dados específicos.

Foram analisados três datasets, correspondentes às ações mencionadas anteriormente (GOLL4 (Gol Linhas Aéreas), PETR4 (Petrobras) e MGLU3 (Magazine Luiza). Esses dados foram retirados diretamente do site Yahoo! Finance e foi utilizado o mesmo período de 01/01/2015 a 01/01/2022 para realizar a análise completa sobre eles.

A escolha dos datasets veio pensando em ações que poderiam ter seus valores influenciados de formas diferentes em virtude da pandemia. Começando pela GOL, empresas relacionadas com ramos de viagens tiveram uma influência muito negativa devido ao lockdown e cancelamentos de viagens. A Magazine Luiza representa o outro lado da moeda, uma empresa nova que cresceu muito na pandemia, pois as pessoas tiveram que recorrer a serviços de entregas no lugar da compra em lojas físicas. Por fim, a Petrobras, sendo uma empresa muito grande que opera há anos e não é tão fácil de ter os valores das suas ações derrubadas por acontecimentos, então seria interessante observar o comportamento durante a pandemia.

Cada um dos datasets possui 1738 linhas e 6 colunas, mas apenas os valores de fechamento das ações foram utilizados para a predição, dando um total de 1738 dados analisados.

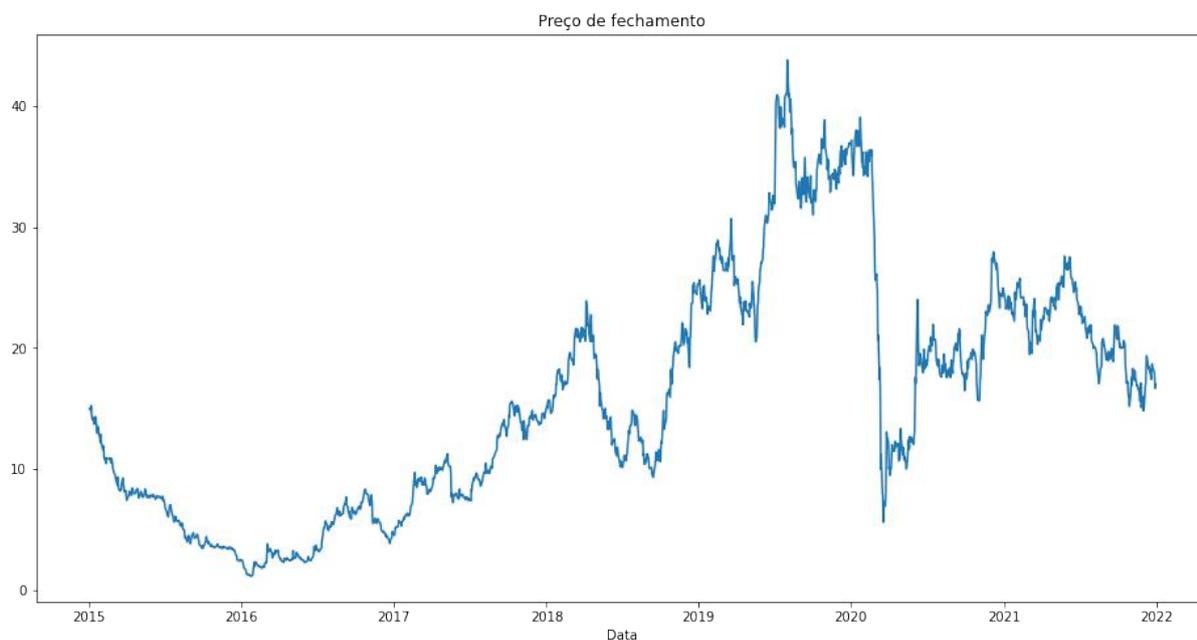
Observando as Figuras 24, 25 e 26, nota-se a queda abrupta no valor das ações no ano de 2020 nos 3 datasets, apesar da queda das ações da Magazine Luiza ser menor que a das outras duas. É interessante notar que essa ação também foi a que se recuperou mais rápido, muito provavelmente por conta dos serviços de compra online, que cresceram muito nesse período.

4.2 CENÁRIOS ANALISADOS

Nesta seção, serão apresentados os diferentes cenários analisados e o objetivo por trás de cada um deles.

Para o desenvolvimento, foram considerados 3 cenários distintos a serem analisados em cada modelo. O primeiro é um cenário sem o período da pandemia, com o intuito de

Figura 24 – Preços de fechamentos da ação GOLL4



Fonte: Elaborado pelos autores

analisar a acurácia dos modelos sem a interferência do problema ocorrido nesse período. Para isso, o dataset foi restringido para o período entre 01/01/2015 a 31/12/2019, sendo utilizado até o dia 31/12/2018 para o treinamento, dando 995 dias. O período restante, até o dia 31/12/2019, foi utilizado para testes, totalizando 248 dias úteis.

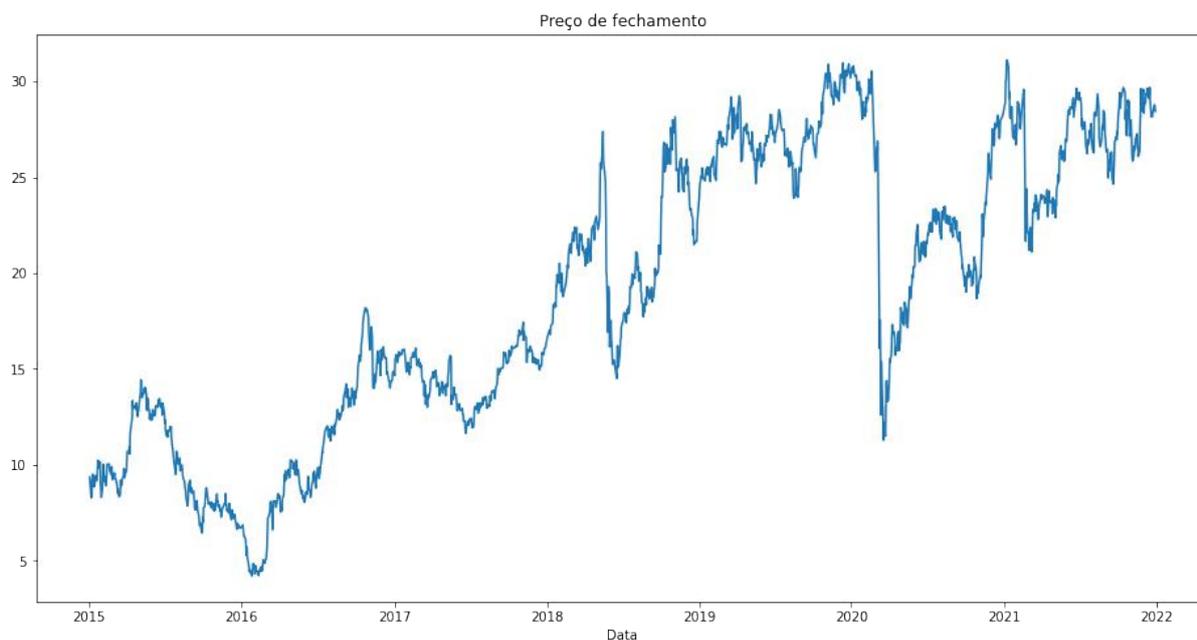
O segundo cenário foi definido visando analisar se a entrada do cenário instável mundial influenciou negativamente no resultado dos modelos. Aqui, o dataset passou a contar com o período de pandemia, sendo utilizado até o dia 31/12/2020 para o treinamento, dando 1491 dias. Para a validação e testes, foram utilizados 247 dias, correspondentes ao período de 01/01/2021 a 31/12/2021.

O terceiro cenário, é um caso onde o dataset é utilizado sem restrições, a fim de verificar se os modelos treinados em períodos estáveis geraram bons resultados para validar períodos de incerteza. Para o treinamento foi utilizado apenas o período pré-pandemia, ou seja, até o dia 31/12/2019, dando 1243 dias. Para a validação e testes foi utilizado o ano inteiro de 2020, dando 248 dias.

4.3 TRATAMENTO DOS DADOS

Antes da aplicação do modelo em si, foi necessário realizar um pré-processamento dos dados para a entrada correta no modelo. Observados os gráficos de cada ação, nota-se que os dados oscilam mais em certos intervalos de tempo do que em outros. Para garantir que eles se comportassem em intervalos de valores semelhantes ao longo do tempo e que

Figura 25 – Preços de fechamentos da ação PETR4



Fonte: Elaborado pelos autores

não possuíssem pesos diferentes, os dados foram normalizados e tratados para a leitura correta de cada modelo.

4.4 CONSTRUÇÃO DOS MODELOS

Nesta seção serão apresentadas as ferramentas utilizadas em cada um dos modelos. De maneira geral, os 3 modelos utilizaram a linguagem Python, na sua versão 3.10. Também foi utilizada a biblioteca YFinance 0.2.18 para retornar os conjuntos de dados diretamente do site do Yahoo.

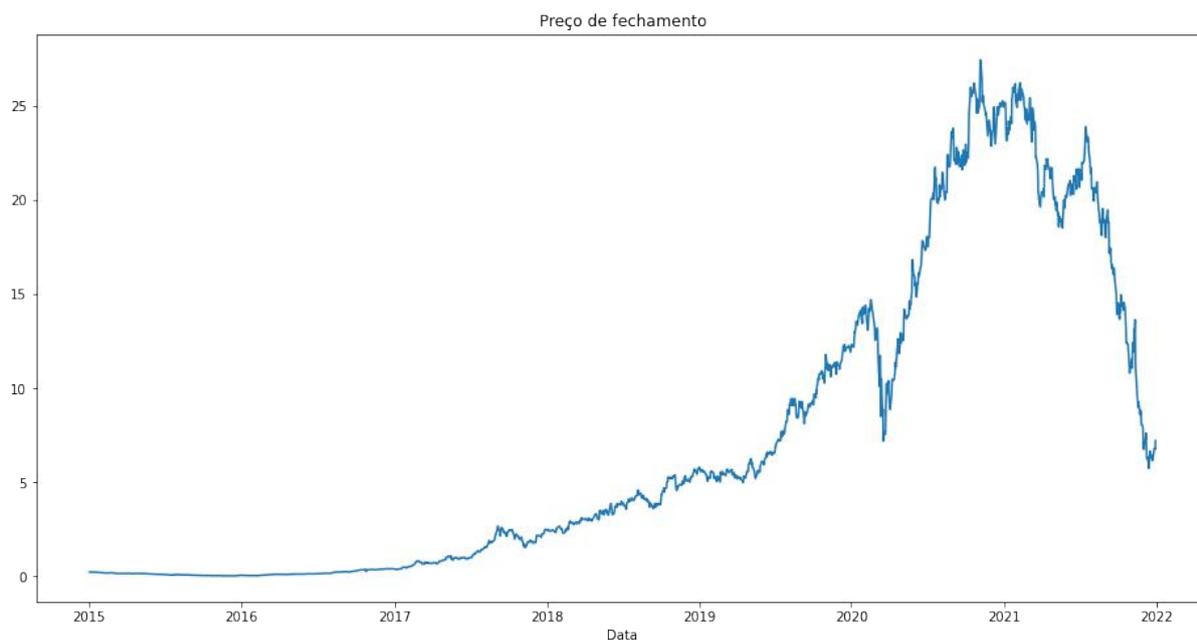
4.4.1 Monte Carlo

Para desenvolver e testar esse modelo usamos bibliotecas para facilitar a manipulação e tratamento dos dados (Ex.:Pandas, Numpy...). Foram utilizadas um milhão de simulações por cenário, para cada ativo.

4.4.2 Sarima

Para desenvolver e testar o modelo de séries temporais foi utilizada, além da linguagem Python, a biblioteca pmdarima que possui uma função de autoArima, onde é possível escolher parâmetros para prever valores de uma série temporal com o modelo SARIMA (VEEN, 2016a).

Figura 26 – Preços de fechamentos da ação MGLU3



Fonte: Elaborado pelos autores

4.4.3 LSTM

A linguagem Python é essencial para esse modelo por permitir a abstração de dados e possui uma abundância de bibliotecas que auxiliam em diversos problemas de aprendizado de máquina. Aqui, foram utilizadas algumas bibliotecas conhecidas para processamento, análise e manipulação de dados como NumPy (versão 1.22.4) e Pandas (versão 1.5.3), além do Scikit-learn (versão 1.2.2), muito utilizada em aprendizado de máquina. A principal biblioteca envolvida foi o Keras (versão 2.12.0), principal biblioteca de redes neurais.

Para a construção do modelo LSTM, foi necessário anteriormente definir o número de camadas da rede. Segundo (ECKHARDT, 2018), um número de camadas considerado bom para problemas mais complexos é 2. Como nesse trabalho a quantidade de dados é elevada e tendo sido utilizados diferentes conjuntos de dados com propriedades diferentes, foi adicionada mais uma camada, para garantir que o modelo se comportaria corretamente. Então, ao todo, o modelo ficou com 3 camadas, onde ele é treinado com os conjuntos de treinamento e recebe o conjunto de testes como entrada para realizar a previsão. Na saída do modelo, ele retorna uma lista com os valores previstos para cada dia do conjunto de testes.

Ainda na construção do modelo, alguns parâmetros foram utilizados para definir a forma de aprendizado da rede, sendo previamente definidos conforme a vontade do programador. Para fazer essa escolha, foram realizados diferentes testes variando o valor de cada parâmetro e observando quais obtinham os menores erros. Segue abaixo uma explicação de cada um deles e o que foi escolhido:

a) Número de neurônios

Foram testadas as quantidades de 16, 32, 64, 128 e 256 neurônios, sendo que os valores de 64 e 128 foram os que apresentaram erro mais baixo. Porém, 64 já é uma abundância de neurônios por camada e quanto maior é esse valor, mais demorado o modelo se torna em seu aprendizado. Por conta disso, foi utilizado 64 neurônios por camada.

b) Dropout

Esse parâmetro possui a finalidade de penalizar a rede, de forma que ela não sofra um overfitting, que nada mais é um caso onde o modelo se ajusta muito bem ao conjunto de dados, mas não consegue prever novos valores. O Dropout é uma porcentagem dos dados que serão escolhidos aleatoriamente para receberem o valor 0 (STATHAKIS, 2009). Aqui, os valores variaram entre 20%, 30%, 40% e 50% e o escolhido foi 20%.

c) Otimizador

O otimizador é um algoritmo que atua na rede neural a fim de escolher os valores ótimos de peso para os neurônios, minimizando assim o erro. Os três mais utilizados são o SGD, o Adam e o RMSProp e aqui, o Adam foi o escolhido porque apresentou o melhor resultado.

d) Tamanho do lote

A cada iteração do algoritmo, uma certa quantidade de dados do conjunto de treino foi utilizada. Foram realizados testes com 8, 16, 32 e 64 para o tamanho do lote e o que apresentou o menor erro foi o tamanho 16. Isso quer dizer que, na primeira iteração, as 16 primeiras amostras foram utilizadas para o treinamento, depois as próximas 16 e assim por diante.

e) Épocas

Uma época representa um momento em que o dataset de treino completo passa pela rede neural. Para valores muito altos o modelo se adaptou a esses dados, ocorrendo um overfitting, enquanto para valores muito baixos o modelo não se ajustou adequadamente. O número de épocas testado começou em 50 e foi aumentando em 50 em cada teste. Para o valor de 150, já começou a aparecer indícios de overfitting, além de aumentar a lentidão para treinar o modelo, o que levou a escolha de 100 épocas.

f) Tamanho da janela

O tamanho da janela diz a quantidade de dias passados que o modelo utiliza para prever o próximo dia. Se o tamanho escolhido for 50, o modelo utiliza dados de 50 dias para prever o 51º dia. Neste trabalho foram testadas janelas de 10 a 80

dias e quanto mais dias menor era o erro apresentado, porém isso implica em uma porcentagem maior do conjunto de teste que não é previsto, o que não é o desejado aqui. Sendo assim, a escolha foi feita baseando-se no menor número de dias possível e que apresentava um erro baixo, e o valor escolhido foi o de 20 dias passados.

5 ANÁLISE DOS RESULTADOS

Neste capítulo serão apresentados os resultados obtidos por cada modelo e cenário, assim como as análises para esses resultados obtidos.

5.1 MONTE CARLO

Agora serão apresentados os resultados obtidos apenas para o modelo considerando o algoritmo de Monte Carlo.

5.1.1 GOL

O modelo estatístico de Monte Carlo apresentou um bom desempenho para indicar tendências nos preços das ações da GOL num panorama. No primeiro cenário tivemos uma acurácia de 88.3% e um MAE (Mean Absolute Error) de 2,72 reais. Pela análise gráfica, é possível ver um bom acompanhamento das previsões. Já no segundo cenário, tivemos uma queda de resultado. 86.1% de acurácia e um MAE de 1,76 reais. Apesar da redução do erro médio, fica nítido, a partir da análise gráfica, que incluindo o período de pandemia no treino, o algoritmo não traz um bom resultado em previsões posteriores ao período. No terceiro cenário, com 90.8% de acurácia e um MAE de 3,06 reais, observamos que, apesar da maior acurácia, o modelo não foi satisfatório em prever a variação de preço devido à pandemia, aumentando o erro médio, porém conseguiu retomar as previsões consistentemente após o período de pandemia.

Figura 27 – Previsões GOL — Cenário 1



Figura 28 – Previsões GOL — Cenário 2



Figura 29 – Previsões GOL — Cenário 3



5.1.2 Petrobras

Paras as ações da Petrobras o modelo apresentou um desempenho parecido com o da GOL, mas com erros menores. No primeiro cenário tivemos uma acurácia de 80.5% e um MAE de 1,37 reais. Já no segundo cenário, tivemos um aumento no resultado. 86.9% de acurácia e um MAE de 1,31 reais, mostrando um comportamento similar ao executado nas ações da GOL. No terceiro cenário, foram obtidos 91.1% de acurácia e um MAE de 1,72 reais, sendo visível graficamente o não acompanhamento da variação de preço em consequência da pandemia.

Figura 30 – Previsões Petrobras — Cenário 1

Predições PETROBRAS - Cenário 1



Figura 31 – Previsões Petrobras — Cenário 2

Predições PETROBRAS - Cenário 2

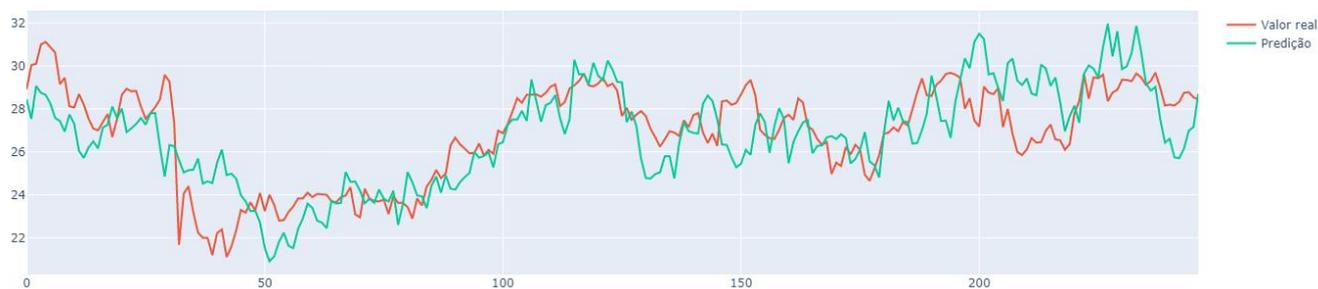
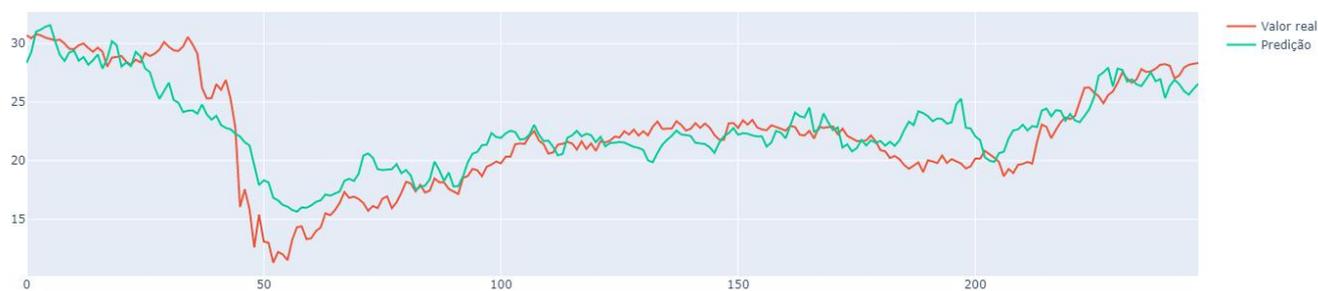


Figura 32 – Previsões Petrobras — Cenário 3

Predições PETROBRAS - Cenário 3



5.1.3 Magazine Luiza

Paras as ações da Magalu o modelo apresentou um comportamento mais intuitivo comparado às outras ações. No primeiro cenário tivemos uma acurácia de 87.8% e um MAE de 0,89 reais. Já no segundo cenário, tivemos uma queda, ao contrário da GOL e da Petrobras. 81.4% de acurácia e um MAE de 3,80 reais. O erro médio sobe bastante. No terceiro cenário, com 87.8% de acurácia e um MAE de 0,89 reais, evidenciando um acompanhamento ruim em relação à pandemia.

Figura 33 – Magalu — Cenário 1



Figura 34 – Magalu — Cenário 2



Figura 35 – Magalu — Cenário 3



5.2 SARIMA

Segundo a Teoria dos Mercados Eficientes de Fama (1970), que relata que os preços do mercado sempre refletem todas as informações existentes, indicando que o preço das ações dificilmente seguirá um padrão definido para se identificar sazonalidades, pois seus preços podem ser influenciados por muitos fatores externos, enquanto as análises em ambientes controlados podem ter um fator sazonal melhor para o modelo de séries temporais. Exemplos destes tipos de análise seriam vendas de algum produto por uma empresa, dados climáticos, etc.

Porém, mesmo com esta questão, já que serão feitas observações intra-anuais para a série temporal, é possível observar alguns comportamentos sazonais como aumento das ações da GOL em finais de ano, onde as pessoas costumam viajar mais. Com isso, a utilização da sazonalidade na análise garantiu resultados melhores que o modelo ARIMA.

5.2.1 GOL

Olhando a figura 36, é possível observar que as previsões ficaram piores no momento que ocorre a pandemia, o que era esperado, visto que seria impossível prever uma pandemia global. Porém, em seguida, as previsões durante a pandemia chegam perto da acurácia das previsões anteriores, mostrando que o método consegue se adaptar depois de um comportamento fora do normal. No primeiro e terceiro casos as previsões conseguem acertar a tendência do preço de fechamento, provavelmente por serem períodos com situações mais controladas.

5.2.2 Petrobras

Através da figura 37 é possível observar que as previsões para este ativo seguem um padrão semelhante ao anterior, diferenciando nas previsões antes da pandemia, onde foi

Figura 36 – Previsões GOL Modelo Sarima



prevista uma queda nos valores, porém o ativo seguiu subindo. O tal equívoco na tendência pode ter sido influenciado pela atitude da Petrobras em aumentar o valor da gasolina refinada no início de 2019, conforme descrito na matéria do G1 Fonte: (G1, 2019)

5.2.3 Magalu

Diferentemente dos ativos anteriores, o Magazine Luiza conseguiu tirar um enorme proveito da pandemia, pois as pessoas não poderiam mais ficar indo às lojas físicas efetuar suas compras. Além disso, a empresa é teoricamente nova, só começando a ter um crescimento em seu valor por volta de 2018, fazendo com que boa parte do treinamento do modelo seja sobre uma série temporal que não variou (começando o treinamento no meio de 2017 iria gerar previsões com maior assertividade).

Figura 37 – Previsões Petrobras Modelo Sarima



A figura 38 evidencia todas as previsões feitas e é possível perceber que o modelo não esperava um crescimento tão grande da empresa. Entre 2017 e 2021 o Magazine Luiza foi uma das empresas que mais cresceu no mundo, principalmente no início da pandemia. Porém, em 2021, conforme o portal de notícias EXAME “O contínuo aumento da inflação, a elevação da taxa básica de juros a patamares não vistos nos últimos quatro anos e o desemprego ainda alto no País desferiram um duro golpe na companhia” (EXAME, 2022), ocasionando uma queda inesperada de 75% das ações da empresa. Este tipo de acontecimento não pode ser previsto analisando apenas valores antigos do ativo, precisaria também de uma análise global sobre inflação, desemprego no país, etc., e isto não é coberto pelo algoritmo de séries temporais, como foi descrito nos capítulos anteriores.

Com isso, ocorreu uma grande inconsistência nas previsões de 2021 da Magalu, conseguindo ter uma assertividade menor que no início da pandemia, mostrando que não é

necessária uma pandemia global para influenciar um ativo inesperadamente e evidenciando o risco que sempre existirá em previsão de valores de ações utilizando séries temporais.

Figura 38 – Previsões Magazine Luiza Modelo Sarima



5.3 REDES NEURAIAS

Apenas revisitando o que foi definido no capítulo anterior, o modelo foi utilizado da seguinte forma para os três conjuntos de dados: 3 camadas com 64 neurônios cada, 20% de dropout, otimizador Adam, lote de tamanho 16, 100 épocas e janela de tamanho 20. Como pode ser visto, as previsões se aproximaram bastante dos valores reais e para a diferença poder ser vista mais facilmente, foi decidido apresentar apenas o conjunto de teste e a previsão nos gráficos seguintes.

5.3.1 GOL

Os resultados encontrados para o dataset da Gol utilizando redes neurais podem ser vistos na Figura 39. Para o primeiro experimento, obteve-se um MAE de 0,21 e é visto que o modelo previu muito bem os valores até o pico que teve após julho de 2019, onde ele passa a prever corretamente os altos e baixos das ações, porém os valores previstos ficaram um pouco abaixo dos reais. Para a segunda previsão, onde os dados de treino não consideraram a pandemia, observa-se uma variação maior entre os valores previstos e os reais no começo da previsão, onde o MAE encontrado foi de 0,26, porém após julho de 2020, os gráficos passam a se sobrepor quase que perfeitamente. Algo parecido ocorre com o terceiro experimento realizado, onde os gráficos passam a se sobrepor na parte final do conjunto, porém aqui o MAE encontrado foi de 0.17, sendo o menor de todos.

Figura 39 – Previsões GOL para o LSTM

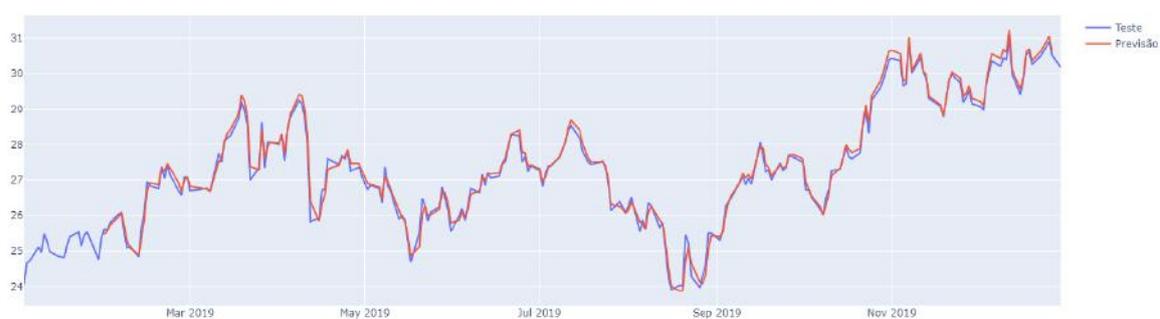


5.3.2 Petrobras

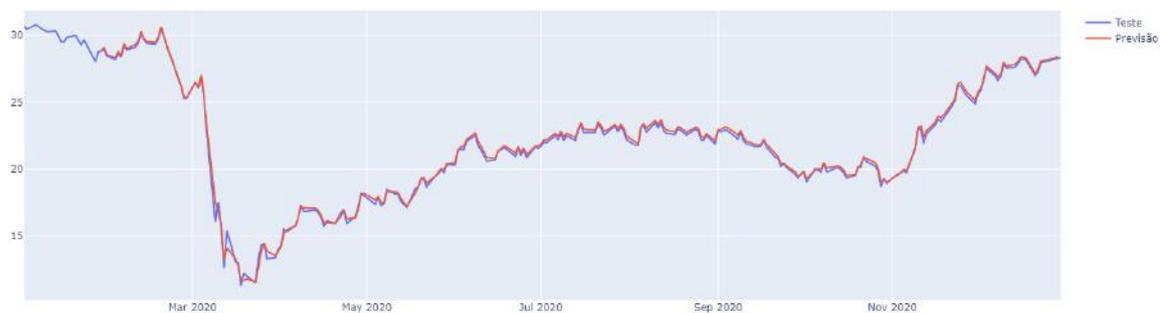
Para as ações da Petrobras, a previsão é visualizada na Figura 40. Aqui, os gráficos não ficaram perfeitamente sobrepostos como no dataset da Gol, apesar do modelo prever corretamente os altos e baixos e alguns períodos do conjunto, mas os valores de erro foram os mais baixos no geral. Isso provavelmente pode ter acontecido, pois o conjunto da Gol foi o escolhido para realizar os testes durante a escolha dos parâmetros, o que pode ter resultado em parâmetros melhores para evitar problemas como overfitting naquele dataset. Os valores encontrados de MAE para esse dataset foram: 0,14 para o primeiro experimento, 0,16 para o segundo e 0,13 no terceiro.

Figura 40 – Previsões Petrobras para o LSTM

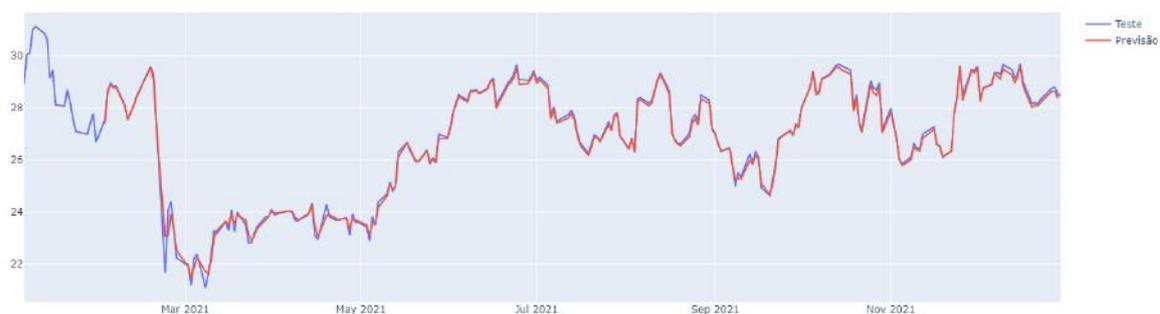
Histórico do preço das ações da Petrobras antes da pandemia



Histórico do preço das ações da Petrobras no início da pandemia



Histórico do preço das ações da Petrobras durante a pandemia



5.3.3 Magazine Luiza

Para as ações da Magazine Luiza, a previsão é visualizada na Figura 41. Para esse conjunto de dados, ocorre o mesmo problema do conjunto da Petrobras, onde os gráficos não ficaram tão sobrepostos. Aqui é onde os valores de MAE ficaram mais distantes entre os experimentos, provavelmente por essa ser uma empresa relativamente nova se comparada com as outras duas e assim, há um número menor de dados e de variações nas ações para serem aprendidos pela rede, então qualquer mudança nos dados causa um impacto significativo. Ainda assim, a previsão se mostrou bem próxima da realidade, demonstrando o quão poderoso é esse modelo, principalmente para o primeiro experimento, que obteve o menor MAE entre os três datasets. O MAE encontrado aqui foi de 0,10 no primeiro experimento, 0,18 no segundo e 0,25 para o último.

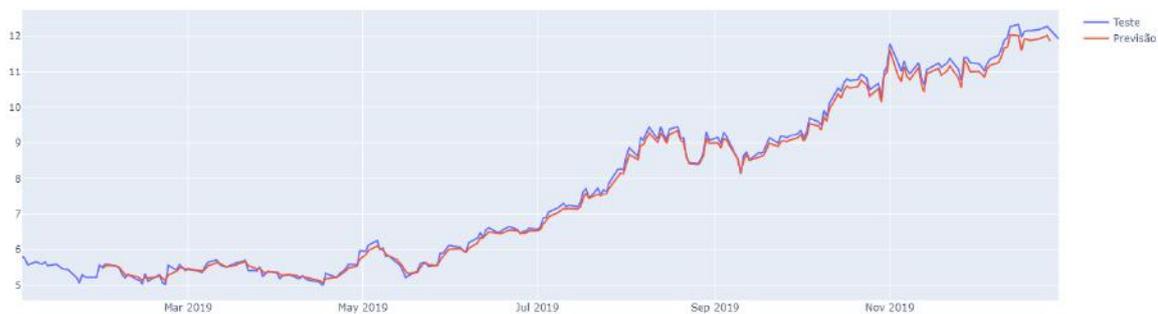
5.3.4 Análise das previsões

Como pode ser visto, o modelo é extremamente eficaz em prever altos e baixos, assim como a tendência que as ações irão tomar. Isso ocorre devido à retropropagação utilizada pela rede. Para uma janela de tamanho 20, a cada iteração da rede, ela utiliza 20 dias do conjunto de treinamento e guarda o valor do 21^o dia como a saída esperada. Quando a previsão é realizada, ela pega 20 dias do conjunto de testes e prevê o 21^o dia utilizando a saída esperada vista anteriormente, porém ela compara essa previsão com o valor real do 21^o dia no conjunto de teste, retornando como saída final um valor próximo do real, observável nos resultados obtidos.

Quando o valor da previsão está distante do valor real no gráfico, isso quer dizer que o modelo tentou prever aquele valor, mas não foi tão eficaz, porém, comparando com o valor real, ele se ajustou para retornar a melhor saída possível e, dessa forma, o resultado da previsão ficou próximo do valor encontrado no conjunto de testes. Por conta disso, o modelo se adaptou muito bem aos 3 experimentos realizados, realizando uma boa previsão, mesmo no caso em que foi necessário prever uma descida tão grande como ocorrido durante a pandemia.

Figura 41 – Previsões Magalu para o LSTM

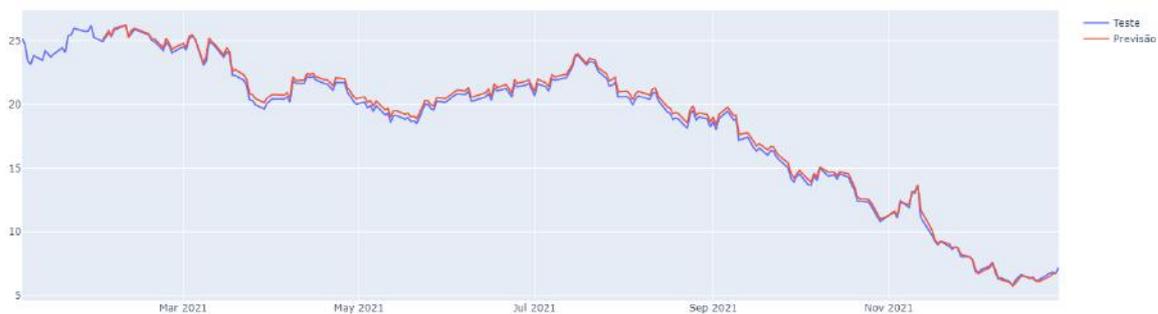
Histórico do preço das ações da Magalu antes da pandemia



Histórico do preço das ações da Magalu no início da pandemia



Histórico do preço das ações da Magalu durante a pandemia



6 CONCLUSÃO

Segundo os resultados, podemos concluir que a sazonalidade exerce uma influência significativa nos resultados das previsões financeiras. A presença de padrões sazonais nos dados, como variações regulares em determinados períodos, pode afetar a precisão das previsões e a capacidade de capturar tendências de longo prazo. Ao realizar previsões financeiras, é crucial considerar e modelar adequadamente a sazonalidade presente nos dados. Ignorar a influência sazonal pode levar a previsões imprecisas e subestimar ou superestimar os valores futuros. Por exemplo, se um modelo não considerar a sazonalidade nas vendas de produtos sazonais, como sorvetes no verão, poderá não conseguir prever corretamente o aumento da demanda durante essa estação, resultando em decisões de negócios ineficientes.

6.1 CONCLUSÃO DOS MODELOS DE PREVISÃO PARA OS ATIVOS E PERÍODOS ESCOLHIDOS

Primeiramente, foi realizada uma conclusão dos 3 modelos separadamente, a fim de mostrar as considerações para cada um deles individualmente.

6.1.1 Conclusão Monte Carlo

Com a análise de todos os cenários para as 3 ações, observamos que o desempenho, no geral, teve um bom resultado. Mas quando entramos no cenário de predição, a variação negativa nos preços devido à pandemia e as simulações não conseguem identificar a tendência de queda, aumentando consideravelmente o erro médio no valor real das ações.

Caso fosse utilizada a técnica para aplicações reais no mercado financeiro, o usuário ficaria refém de alterações macro-econômicas para obter bons resultados no investimentos.

6.1.2 Conclusão SARIMA

Analisando todas as previsões feitas pelo modelo de séries temporais, podemos concluir que ele não é muito bem aproveitado em previsão de valores de ações, pois o modelo não considera o maior fator de influência que são acontecimentos externos que ocorreram paralelamente. A previsão em séries temporais está sempre prevendo de acordo com valores anteriores da série.

Portanto, o modelo tem seu uso mais assertivo em previsões de séries menos influenciadas por meios externos, como previsões do número de focos de calor realizada por Mariano Martinez Espinosa em (ESPINOSA, 2010). Mesmo existindo um fato externo como aquecimento global, ele não irá influenciar de forma rápida e pode ser estudado utili-

zando os dados anteriores, possibilitando um estudo sobre a proporção do crescimento dos focos de calor nos últimos anos. Este tipo de série temporal estudada possui tendências mais definidas e o fator sazonalidade pode ser muito bem observado ao decorrer da série.

6.1.3 Conclusão LSTM

Após a realização das previsões, foi vista a capacidade imensa de aprendizado e previsão das redes neurais. A rede LSTM se mostrou extremamente eficiente em prever os momentos de altos e baixos das variações da bolsa, como a previsão de descida das ações durante o período de pandemia. Mesmo para as ações que não foram consideradas para os testes de escolha de parâmetros da rede, ela conseguiu prever a tendência de subida e descida desses conjuntos de dados.

Sendo assim, se fossem realizados diferentes testes e escolhidos diferentes parâmetros para cada dataset, a rede se adaptaria melhor ainda e os valores previstos para cada um se aproximariam ainda mais da realidade. Um problema dessa rede é que ela se comporta muito bem prevendo os valores e comparando com valores já conhecidos, porém para prever o futuro, ou seja, para valores além do conjunto de testes, essa mesma abordagem não é possível, pois não há valores ainda para serem comparados na previsão. Dessa forma, seria necessário outro tratamento de forma que a rede consiga prever um valor no futuro e utilize esse mesmo valor para continuar as próximas previsões a cada iteração.

6.2 CONCLUSÃO COMPARANDO OS 3 MODELOS

Falando sobre o trabalho, agora olhando para os três modelos utilizados, de forma geral. Os três se encaixam em áreas diferentes da computação visando a predição de valores. Embora essa distinção, os três possuem a mesma estratégia de utilizar apenas o fator de variação no tempo da série temporal em sua análise e treinamento. Isso, explicitamente, nos diz que fatores que não sejam as datas, não influenciam na tomada de decisão para geração do resultado. Analisando os resultados, vemos que as redes neurais conseguem desempenhar um controle bem melhor dos fatores externos. A possibilidade de configuração de parâmetros, mais especificamente o batch size, no LSTM garante que alterações recentes sejam utilizadas como influenciador de tendência e predição. Com isso, o modelo de deep learning não só lidou melhor com a variação na pandemia como demonstrou forte resiliência na recuperação pós-pandemia e no ajuste dos valores futuros.

6.3 CONCLUSÃO SOBRE DESENVOLVIMENTO EM GRUPO

Geralmente, os trabalhos de conclusão de curso são realizados individualmente. Entretanto, tivemos a oportunidade de desenvolver nosso projeto em grupo, composto por três pessoas, e podemos afirmar que essa abordagem favoreceu enormemente o seu desen-

volvimento. Optamos por dividir o trabalho em três partes, onde cada integrante abordou um modelo específico, e os outros dois membros leram e revisaram o conteúdo produzido, sem possuir conhecimento prévio sobre o tema. O objetivo era garantir que uma pessoa sem familiaridade com o assunto compreendesse o funcionamento de cada modelo ao ler o nosso artigo.

Essa forma de trabalho se baseou na responsabilização individual de cada membro pelo seu tópico, evitando situações em que um membro transferisse a dificuldade para o outro. Acreditamos que essa abordagem promoveu um ambiente de comprometimento e colaboração, onde todos assumiram a responsabilidade por suas partes no projeto.

Ao final, podemos concluir que essa experiência foi extremamente proveitosa para todos os envolvidos, e os resultados alcançados superaram as nossas expectativas iniciais. Acreditamos que essa abordagem de trabalho em grupo pode servir como modelo para futuros trabalhos de conclusão de curso, pois uma maior troca de conhecimento estimula o senso de responsabilidade individual e contribui para um produto final mais abrangente e compreensível.

REFERÊNCIAS

- ANBIMA. **Raio X do Investidor Brasileiro**. 2021. Disponível em: https://www.anbima.com.br/pt_br/especial/raio-x-do-investidor-2021.htm. Acesso em: 01 jul. 2022.
- ECKHARDT, K. **Choosing the right Hyperparameters for a simple LSTM using Keras**. 2018. Disponível em: <https://towardsdatascience.com/choosing-the-right-hyperparameters-for-a-simple-lstm-using-keras-f8e9ed76f046>. Acesso em: 19 jul. 2023.
- ESPINOSA, M. M. Uso do modelo sarima na previsão do número de focos de calor para os meses de junho a outubro no estado de mato grosso. v. 1, 2010.
- EXAME. **Magalu2021**. 2022. Disponível em: <https://exame.com/invest/onde-investir/por-que-as-acoes-do-magazine-luiza-cairam-75-em-12-meses/>. Acesso em: 30 nov. 2022.
- FAVA, V. L. Análise de séries de tempo. manual de econometria : nível intermediário. 2000.
- FIESP, V. **GraficoExemploVenda**. 2019. Disponível em: <https://www.fiesp.com.br/observatoriodaconstrucao/noticias/comercio-de-materiais-de-construcao-registra-crescimento-de-43-em-2019/>. Acesso em: 12 jul. 2022.
- G1. **Petrobras2019**. 2019. Disponível em: <https://g1.globo.com/economia/noticia/2019/01/16/petrobras-eleva-em-25-preco-medio-da-gasolina-nas-refinarias.ghtml>. Acesso em: 30 nov. 2022.
- GIRIOLI, L. S.; RIBEIRO, E. M. S. Utilização de redes neurais artificiais para análise técnica no mercado de ações: Estudo dos Índices ibovespa e dow jones industrial average (djia). **Revista EPeQ/Fafibe**, São Paulo, v. 1, n. 1, p. 150–155, 2009.
- HAYKIN, S. **Redes Neurais: Princípios e Prática**. Tradução de: P. M. Engels. 2. ed. Porto Alegre: Bookman, 2001. 898 p.
- HAYKIN, S. **Neural Networks and Learning Machines**. 3. ed. New Jersey: Pearson, 2009.
- KARPATHY, A. **The Unreasonable Effectiveness of Recurrent Neural Networks**. 2015. Disponível em: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>. Acesso em: 03 ago. 2022.
- LA, F. L. Como as redes neurais aprendem? **MSDN Magazine Issues**, v. 34, n. 4, 2019.
- MAMANDIPOOR, B. et al. Monitoring and detecting faults in wastewater treatment plants using deep learning. **Environmental Monitoring and Assessment**, v. 192, n. 2, 2020.

MOGHADDAM, A. H.; MOGHADDAM, M. H.; ESFANDYARI, M. Stock market index prediction using artificial neural network. **Journal of Economics, Finance and Administrative Science**, v. 21, n. 41, p. 89–93, 2016.

OLAH, C. **Understanding LSTM Networks**. 2015. Disponível em: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acesso em: 21 set. 2022.

RABELO, L. **Princípios básicos para criar previsões de Séries Temporais**. 2019. Disponível em: <https://bityli.com/dMDgRRp>. Acesso em: 12 jul. 2022.

SARDINHA, V. **Neurônio**. 2019. Disponível em: <https://mundoeducacao.uol.com.br/biologia/neuronios.htm>. Acesso em: 02 jul. 2022.

SOUZA, L. V. de. Programação genética e combinação de preditores para previsão de séries temporais. v. 1, p. 18–21, 2006.

STATHAKIS, D. How many hidden layers and nodes? **International Journal of Remote Sensing**, Taylor & Francis, v. 30, n. 8, p. 2133–2147, 2009. Disponível em: <https://doi.org/10.1080/01431160802549278>.

UNICAMP. **Distribuição**. 2019. Disponível em: <http://me414-unicamp.github.io/aulas/slides/parte13/parte13.html#10>. Acesso em: 01 jul. 2023.

VASCO, L. P. **Um Estudo de Redes Neurais Recorrentes no Contexto de Previsões no Mercado Financeiro**. Tese (Graduação) — Universidade Federal de São Carlos, São Carlos, Brasil, Dezembro 2020.

VEEN, F. van. **Documentação pmdarima**. 2016. Disponível em: <http://alkaline-ml.com/pmdarima/>. Acesso em: 02 jul. 2022.

VEEN, F. van. **The Neural Network Zoo**. 2016. Disponível em: <https://www.asimovinstitute.org/neural-network-zoo/>. Acesso em: 02 jul. 2022.

VENUGOPAL, P.; T., V. State-of-health estimation of li-ion batteries in electric vehicle using indrnn under variable load condition. **Energies**, v. 12, n. 22, 2009.

VIDAL, F. et al. Projeto e implementação de uma rede neural artificial para detecção do mal-posicionamento rotacional de dedos em dispositivos de captura de impressões digitais multivista sem toque. In: **XI Brazilian Symposium on Information System**. Goiânia, GO: [s.n.], 2015. p. 211–218.

WLODZISLAW, D.; JANKOWSKI, N. Survey of neural transfer functions. **Neural Computing Surveys**, v. 2, p. 163–212, 1999.

ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. **International Journal of Forecasting**, v. 14, n. 1, p. 35–62, 1998.