

Coelho, Filipe Ferreira

Machine learning e análise técnica como ferramentas para construção de portfólios de renda variável no mercado brasileiro / Filipe Ferreira Coelho. – 2020.

89 f.

Dissertação (mestrado) - Fundação Getulio Vargas, Escola Brasileira de Economia e Finanças.

Orientador: Gustavo Silva Araújo

Inclui bibliografia.

1. Aprendizado do computador. 2. Redes neurais (Computação). 3. Mercado de capitais - Brasil. 4. Investimentos. 5. Python (Linguagem de programação de computador). I. Araújo, Gustavo Silva. II. Fundação Getulio Vargas. Escola Brasileira de Economia e Finanças. III. Título.

CDD – 621.39

FILIFE FERREIRA COELHO

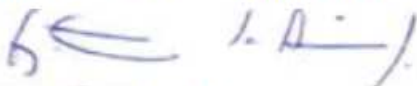
"MACHINE LEARNING E ANÁLISE TÉCNICA COMO FERRAMENTAS PARA CONSTRUÇÃO DE PORTFÓLIOS DE RENDA VARIÁVEL NO MERCADO BRASILEIRO".

Dissertação apresentado (a) ao Curso de Mestrado Profissional em Economia Empresarial e Finanças do(a) EPGE Escola Brasileira de Economia e Finanças - FGV EPGE para obtenção do grau de Mestre(a) em Economia.

Data da defesa: 15/10/2020

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Presidente da Comissão Examinadora: Prof^o Gustavo Silva Araujo



Gustavo Silva Araujo
Orientador



Rafael Martins de Souza
Membro Interno



Wagner Piazza Gaglianone
Membro Externo

Nos termos da Lei nº 13.979 de 06/02/20 - DOU nº 27 de 07/02/20 e Portaria MEC nº 544 de 16/06/20 - DOU nº 114 de 17/06/20 que dispõem sobre a suspensão temporária das atividades acadêmicas presenciais e a utilização de recursos tecnológicos face ao COVID-19, as apresentações das defesas de Tese e Dissertação, de forma excepcional, serão realizadas de forma remota e síncrona, incluindo-se nessa modalidade membros da banca e discente.



Ricardo de Oliveira Cavalcanti
Coordenador



Antonio de Araujo Freitas Junior
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação FGV

Antonio Freitas, PhD
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação
Fundação Getúlio Vargas

Instrução Normativa nº 01/19, de 09/07/19 - Pró-Reitoria FGV

Em caso de participação de Membro(s) da Banca Examinadora de forma não-presencial*, o Presidente da Comissão Examinadora assinará o documento como representante legal, delegado por esta I.N.

*Skype, Videoconferência, Apps de vídeo etc

**FUNDAÇÃO GETÚLIO VARGAS
ESCOLA BRASILEIRA DE ECONOMIA E FINANÇAS**

FILIPPE FERREIRA COELHO

**MACHINE LEARNING E ANÁLISE TÉCNICA COMO FERRAMENTAS
PARA CONSTRUÇÃO DE PORTFÓLIOS DE RENDA VARIÁVEL NO MERCADO
BRASILEIRO**

Rio de Janeiro

2020

FILIPPE FERREIRA COELHO

**MACHINE LEARNING E ANÁLISE TÉCNICA COMO FERRAMENTAS
PARA CONSTRUÇÃO DE PORTFÓLIOS DE RENDA VARIÁVEL NO MERCADO
BRASILEIRO**

Dissertação para obtenção do grau de mestre apresentada à Escola
Brasileira de Economia e Finanças

Área de concentração: Finanças

Orientador: Gustavo Silva Araujo

Rio de Janeiro

2020

AGRADECIMENTOS

Agradeço primeiramente a Deus pela oportunidade de ter cursado esse mestrado, que foi um grande diferencial na minha vida acadêmica.

Agradeço aos meus pais, Claudio e Fátima, meus pilares na vida e por todo sacrifício e dedicação aos meus estudos e à minha felicidade. São o motivo e o que me dão forças todos os dias para continuar.

Agradeço à Fundação Getúlio Vargas (FGV) pelo suporte oferecido durante todo mestrado e aos professores da EPGE do Mestrado Profissional em Economia e Finanças por todo o conhecimento disseminado dentro e fora da sala de aula. Em especial ao professor e orientador Gustavo Silva Araujo e aos colegas e amigos Felipe, Flávia e Jessica que me deram todo o apoio e parceria.

Agradeço ao Banco do Brasil e à BB DTVM por apoiar com o tempo necessário para essa pesquisa. Com destaque para aos colegas do CESUP, Modelagem de Riscos e Gestão Offshore por toda ajuda, paciência e conhecimento proporcionados durante o tempo de mestrado. Destes um carinho especial pelos colegas Andrea, Rafael e Rose por confiar no meu trabalho e potencial. Aos colegas Flavio, Gustavo e João Ricardo por me receberem tão bem e me auxiliarem no novo desafio que já dá frutos do seu sucesso.

Agradeço aos meus irmãos de outros pais Guilherme, Juliana, Matheus e Patrick obrigado pelo carinho e paciência ao longo de todos os anos de amizade, ajudando em quase todas as etapas da minha vida e sempre prontos para um conselho ou uma palavra amiga sempre que precisei.

RESUMO

Em um cenário de crescimento global da quantidade de fundos quantitativos verifica-se uma necessidade de ampliar os estudos voltados a esta temática de fundos no Brasil. A ideia para o desenvolvimento do modelo deste trabalho iniciou-se com a percepção do aumento da competitividade da indústria de distribuição de investimentos, cada dia provendo soluções melhores e mais acessíveis em termos de custos e diversificação a diversos seguimentos de clientes. Com o advento de tecnologias acessíveis para previsão a partir de séries de dados, um novo leque de possibilidades para o mundo de finanças foi aberto. Com o objetivo de construir portfólios de ativos de renda variável, foi utilizado um algoritmo de *machine learning*, chamado *Long-Short Term Memory* (LSTM), e indicadores de análise técnica, através da linguagem Python. Deste modo, o estudo propõe a obtenção de carteiras com uma melhor relação risco-retorno em relação ao seu *benchmark*. Para o estudo foi utilizada uma amostra composta de 50 ações do Ibovespa, com informações de preços entre 2009 e 2018, construiu-se indicadores de análise técnica a partir destes dados e, com essas informações geradas, utilizou-se o algoritmo LSTM para prever o retorno esperado, a partir destes retornos a volatilidade, ambas para o período de 2019. Destes resultados foi utilizada a metodologia de alocação proposta por Markowitz para a obtenção do portfólio de melhor índice de Sharpe e o de menor volatilidade. Após a utilização da metodologia descrita, obteve-se seis portfólios, três desses com o melhor índice de Sharpe e três com a menor volatilidade. Adicionalmente, dado o grande estresse vivenciado no mundo causado pela pandemia do COVID-19, com destaque para o mercado financeiro, os mesmos portfólios foram avaliados em simulações que contemplam apenas o período de janeiro de 2020 a setembro de 2020, com o intuito de testar o modelo em condições severas de estresse de mercado.

Palavras-chave: *machine learning*, LSTM, redes neurais, otimização de portfólio, renda variável

ABSTRACT

As the number of quantitative funds increase worldwide, *pari passu*, there is a necessity to expand the research in this area in Brazil. The idea behind the model proposed in this project started with the perception of increase in competition in the industry of financial investments, with companies providing better and more affordable solutions in terms of cost and diversification through several client segments. As new accessible data prediction technologies are launched, a whole new range of possibilities for the financial industry becomes available. Aiming the construction of Brazilian equity portfolios, a machine learning algorithm called Long-Short Term Memory (LSTM) and technical analysis indicators, through Python coding language, were used in this research. Hence, this study proposes obtaining portfolios with the best risk/return ratio comparatively with its benchmark. For this study the information of price and volume, between 2009 to 2018, for a fifty equity stocks sample of Bovespa Index (Ibovespa) were used to build technical analysis indicators and with that information the LSTM algorithm was used to generate the forward predicted returns for the 2019 year. With these results the portfolio optimization proposed by Markowitz was used to obtain the higher Sharpe Ratio and the lowest volatility portfolios. Upon the use of the described methodology six portfolios were created: three with the best Sharpe Ratio and three with the lowest volatility. Additionally, given the great stress experienced in the world caused by COVID-19 pandemic, especially in the financial markets, the same portfolios were tested in simulations that only cover the period from January 2020 to September 2020, in order to test the model in severe stressful conditions.

Keywords: machine learning, LSTM, neural network, portfolio optimization, equity

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura simplificada de uma Rede Neural Recorrente.....	33
Figura 2 - Representação de uma célula LSTM.	35
Figura 3 – Ilustração da divisão da base de dados	51
Figura 4 – Ilustração do modelo de previsão e seu funcionamento.....	53

LISTA DE TABELAS

Tabela 1 - Ações escolhidas para seleção de portfólio.	49
Tabela 2 - Retornos mensais dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e o Ibovespa, de todos os meses do ano de 2019.	62
Tabela 3 – Decomposição do portfólio Max Sharpe 3d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	64
Tabela 4 - Decomposição do portfólio Max Sharpe 4d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	65
Tabela 5 - Decomposição do portfólio Max Sharpe 5d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	66
Tabela 6 - Decomposição do portfólio Min Vol 3d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	67
Tabela 7 - Decomposição do portfólio Min Vol 4d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	68
Tabela 8 - Decomposição do portfólio Min Vol 5d por percentual alocado em cada setor e ativo do portfólio e do <i>benchmark</i> e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.	69

LISTA DE GRÁFICOS

Gráfico 1 - Gráfico Google Trends: Busca pelo termo “machine learning time series” nos últimos 5 anos. Período: 12/05/2014 a 31/12/2019.....	31
Gráfico 2 - Comparativo entre os retornos acumulados dos portfólios que efetuam a minimização da volatilidade e do Ibovespa entre os períodos de janeiro de 2019 a dezembro de 2019.	55
Gráfico 3: Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe e do Ibovespa entre os períodos de janeiro de 2019 a dezembro de 2019.....	56
Gráfico 4 - Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe e do Ibovespa entre os períodos de janeiro de 2019 a agosto de 2020.	57
Gráfico 5 - Comparativo entre os retornos acumulados dos portfólios que efetuam a minimização da volatilidade e do Ibovespa entre os períodos de janeiro de 2019 a agosto de 2020.	58
Gráfico 6 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre o período de janeiro de 2019 a dezembro de 2019.....	59
Gráfico 7 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre o período de janeiro de 2019 a agosto de 2020.	60

LISTA DE ANEXOS

Anexo 1 - Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe, dos portfólios que minimizam a volatilidade e do Ibovespa, entre as datas de 01/01/2020 a 31/08/2020.....	82
Anexo 2 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre as datas de 01/01/2020 a 31/08/2020.	83
Anexo 3 - Gráfico <i>Drawdown</i> (Perdas). Mínima Volatilidade x Máximo Sharpe x Ibovespa (01/01/2019 – 31/08/2020).....	84
Anexo 4 - Tabela com rentabilidades mensais. Mínima Volatilidade x Máximo Sharpe x Ibovespa (01/2019 – 08/2020).....	85
Anexo 5 - Composição portfólio Mínimo Volatilidade, com previsões de cinco dias.	86
Anexo 6 - Composição portfólio Máximo Sharpe, com previsões para cinco dias	86
Anexo 7 - Composição portfólio Mínimo Volatilidade, com previsões para quatro dias	87
Anexo 8 - Composição portfólio Máximo Sharpe, com previsões para quatro dias	88
Anexo 9 - Composição portfólio Mínimo Volatilidade, com previsões para três dias.	88
Anexo 10 - Composição portfólio Máximo Sharpe, com previsões para três dias.....	89
Anexo 11 - Relação retorno acumulado e volatilidade anualizada dos portfólios e de todas as ações individuais que os compõem os portfólios no período entre janeiro de 2019 a agosto de 2020.....	Erro! Indicador não definido.

Lista de Siglas

3FM	Fama and French Three Factor Model
A/D	Acumulação/Distribuição
BB	Bandas de Bollinger
CAPM	Capital Asset Pricing Model
DAX	German Deutscher Aktien
DJLA	Dow Jones Industrial Average
EQWT	Equally Weighted Portfolio Rebalancing
ETF	Exchange-traded Fund
FTSE	British Financial Times Stock Exchange
IS	Índice de Sharpe
LASSO	Least Absolute Shrinkage and Selection Operator
LSTM	Long Short Term Memory
MACD	Moving Average Convergence Divergence
TME	Teoria do Mercado Eficiente
ML	Machine Learning
MME	Média Móvel Exponencial
MMEP	Média Móvel Exponencialmente Ponderada
MMS	Média Móvel Simples
MVO	Mean Variance Optimization
ROC	Rate of Change
RSI	Relative Strength Index
SELIC	Sistema Especial de Liquidação e Custódia
SVM	Support Vector Machine
TME	Teoria do Mercado Eficiente
TSI	True Strenght Index

SUMÁRIO

Capítulo 1 – INTRODUÇÃO	12
Capítulo 2 – REFERÊNCIAL TEÓRICO	15
2.1 Mercado de Ações	15
2.1.1 Teoria do Mercado Eficiente	16
2.1.2 Análise Técnica	17
2.2 Machine Learning	29
2.2.1 Redes Neurais	33
2.2.2 Long-Short Term Memory (LSTM)	34
2.2.3 Previsões	36
2.3 Construção de Portfólio	39
2.3.1 Índice de Sharpe	41
2.3.2 Modelo de Otimização Média-Variância.....	42
Capítulo 3 METODOLOGIA E AMOSTRA	47
3.1 Seleção da amostra e Base de Dados	47
3.2 Experimento.....	51
Capítulo 4 RESULTADOS.....	54
4.1 Análises dos Resultados – RETORNOS	55
4.2 Análises dos Resultados – RISCO-RETORNO	59
4.3 Análises dos resultados – DECOMPOSIÇÃO DE CARTEIRAS.....	63
Capítulo 5 CONCLUSÃO	72
Referências Bibliográficas.....	73
Capítulo 6 Anexo.....	82

CAPÍTULO 1 – INTRODUÇÃO

A teoria do mercado eficiente (TME) de Fama (1970) nos prediz que os preços dos ativos refletem todas as informações passadas e, no evento de novas informações, há um ajuste no preço para que o ativo seja precificado corretamente.

Entretanto, alguns estudos rejeitam a teoria de mercado eficiente, argumentando que o comportamento dos retornos não é aleatório e que series temporais de finanças, de fato, possuem memória de longo prazo. Evidências empíricas apresentadas por Cervelló-Royo et al. (2015) comprometem a teoria. Os autores desenvolveram um modelo baseado em análise técnica para analisar o desempenho intradiário de três índices no período de 2000 a 2013: American Dow Jones Industrial Average (DJIA), German Deutscher Aktien 30 Index (DAX) e British Financial Times Stock Exchange (FTSE). Conforme De Carvalho et al. (2008), o termo “análise técnica” compreende uma ampla coleção de técnicas de negociações que em sua essência envolvem a busca de padrões recorrentes e previsíveis nos preços das ações. O objetivo da análise, além de fornecer subsídios para a análise dos ativos, é antecipar possíveis movimentos antes que essa informação se torne um consenso para o mercado. Sendo assim, a análise deve possibilitar ao analista se antecipar ao restante do mercado na realização de operações. Esta afirmação contraria a TME, pois esta afirma um agente não consegue alcançar consistentemente retornos superiores à média do mercado. Os resultados de Cervelló-Royo et al. (2015) se mostraram positivos para os três mercados estudados. Chourmouziadis e Chatzoglou (2016) pesquisaram retornos de portfolios montados utilizando análise técnica com uma janela diária de dados da bolsa de valores de Atenas entre o período de 1996 até 2012. Os resultados também se mostraram positivos.

O objetivo do presente trabalho é construir uma carteira exclusivamente de ações, com um baixo custo computacional e operacional, para auxiliar na tomada de decisão e na alocação de carteiras de investimento a partir de índices de análise técnica. Para obter a melhor carteira, são utilizadas as técnicas de alocação de Markowitz. A amostra utilizada contém 50 ações que compõem o Ibovespa, com preços diários no período de 2008 a 2018.

Para o fornecimento de dados ao modelo de Markowitz foram efetuadas previsões para os retornos diários de todas as ações utilizando uma técnica de *machine learning*, com um algoritmo de redes neurais chamado LSTM (*Long Short Term Memory*) e calculados os desvios-padrão e as correlações, a partir dos retornos previstos. O algoritmo foi treinado com dados observados dos ativos (preço de fechamento, abertura, volume) e dados de análise técnica

(índice de força relativa, oscilador estocástico, MACD, entre outros)¹, ou seja, a previsão dos retornos foi efetuada a partir desses dados. A vantagem do uso de redes neurais para previsões a partir de séries temporais é a capacidade da rede aprender a partir dos dados que lhe foram fornecidos e com isso capturar dependências não-lineares e contínuas.

A linguagem Python foi utilizada devido a sua larga usabilidade e a disponibilidade de pacotes que permitem a manipulação e tratamento de dados. Através desta ferramenta as previsões de preços, foram obtidas através de *scripts*, programados em uma plataforma para análise de dados e para trabalhar com *machine learning*, denominada Anaconda².

A construção dos portfólios também foi efetuada através da linguagem Python. O *script* desenvolvido efetua a leitura da base de dados com as previsões efetuadas pelo modelo e assim calcula o desvio-padrão dos retornos de cada ação e a matriz de covariância entre eles.

Cria-se um total de seis carteiras, cada uma contendo um número diferente de ativos e percentuais de alocação. Três destas carteiras representam portfólios elaborados com intuito de minimizar a volatilidade enquanto os outros três têm por objetivo maximizar o índice de Sharpe.³ Para todos os portfólios foram feitas previsões de retorno para três, quatro e cinco dias úteis.

Os resultados mostram que os três portfólios que têm por objetivo minimizar a volatilidade obtiveram um retorno total, no período analisado, semelhante ao *benchmark* (Ibovespa) com uma volatilidade menor. Os três portfólios cujo objetivo é maximizar o índice de Sharpe possuem uma melhor relação risco-retorno, quantificado pelo Índice de Sharpe, quando comparados ao seu *benchmark*.

Este trabalho procurou contribuir com um estudo cujo resultado seja possível ser implementado e utilizado na prática, de uma maneira acessível, por qualquer leitor. Inova-se na pesquisa no que diz respeito à união de três pilares: análise técnica, *machine learning* e otimização de portfólio, com objetivo de se entregar um produto tangível, cujos resultados possam ser verificados na prática no mercado de capitais brasileiro. Apesar da existência de estudos que façam a previsão de séries temporais com *machine learning* e estudos que utilizem análise técnica com otimização de portfólio, não se verifica na literatura vigente nenhum que faça a metodologia proposta neste trabalho. O estudo que mais se assemelha ao presente

¹ MACD é a sigla para Moving Average Convergence Divergence, que, em português, significa Média Móvel Convergente e Divergente.

² Bibliotecas da linguagem Python utilizadas para a elaboração do trabalho foram: Pandas, Numpy, Matplotlib, Scikit-Learn e Keras.

³ Minimizar a volatilidade neste contexto refere-se a minimizar a volatilidade em relação ao *benchmark*, neste caso prático o Ibovespa. Somente ativos de renda variável foram considerados.

trabalho seria o de Obeidat et al. (2018). Os autores também se utilizam de retornos de ativos construídos a partir da rede LSTM em conjunto com otimização de portfólio de Markowitz. Entretanto, os autores se separam deste na quantidade e nos ativos utilizados (11 ativos negociados no mercado de ações americano contra 50 ativos do Ibovespa) e não se utilizaram de nenhum indicador de análise técnica. Contribui-se com o desenvolvimento de portfólios de renda variável acessíveis aos investidores comuns e com processo de gestão de carteira sem nenhuma interferência humana no que tange à tomada de decisões sobre percentuais de alocação ou quaisquer rebalanceamentos de carteira.

O presente estudo adiciona à literatura de investimentos uma abordagem não ainda vista dado o modelo proposto e enriquece os estudos de análise técnica com utilização de algoritmos de *machine learning*. Esta pesquisa adicionou a crescente base de estudos que versam sobre análise técnica e machine learning a abordagem com otimização de portfólio e a utilização do algoritmo LSTM com os 15 índices e indicadores de análise técnica aqui descritos. O estudo se distingue de Obeidat et al. (2016) pelo modelo ser aplicado no Brasil e por ser com indicadores de análise técnica. Portanto, é uma adição inédita à literatura.

A partir deste estudo, conclui-se que é possível utilizar ferramentas modernas, como Python e redes neurais, para elaboração de portfólios de renda variável em conjunto com técnicas de alocação de recurso de Markowitz e indicadores técnicos. Os resultados mais uma vez, como os trabalhos de Barbosa (2007), Boainain (2007) e Chicaroli e Valls Pereira (2008), indicam uma rejeição à teoria de mercado eficiente (TME) que neste caso pode ser relacionado à ineficiência do mercado emergente como o Brasil ou a fatores comportamentais e psicológicos que influenciam a previsão de preços de ativos.

CAPÍTULO 2 – REFERÊNCIAL TEÓRICO

Este capítulo é subdividido em três seções com o propósito de abordar conceitos teóricos importantes para a pesquisa e criar um arcabouço para subsidiar o seu desenvolvimento. As seções versam sobre: Mercado de Ações, *Machine Learning* e Construção de Portfólios.

2.1 MERCADO DE AÇÕES

O surgimento do mercado de ações pode ser datado de mais de cem anos antes do processo de industrialização começar na Europa. O rápido crescimento econômico dos últimos anos deu luz ao desenvolvimento do mercado de ações e fez deste um pilar fundamental para a economia mundial. O moderno sistema econômico tem como um dos seus pilares um mercado de ações que ofereça informações sobre o custo de capital atual de cada empresa, para determinar a magnitude e estrutura das companhias. O mercado também provê informações do fluxo financeiro de investidores, uma vez que os preços de ações são sinais para a demanda de cada mercado.

Mercados de ações no mundo inteiro desenvolveram-se de maneira vertiginosa nos últimos trinta anos. De acordo com De Camargos e Barbosa (2006), podemos perceber vários benefícios deste crescimento como o desenvolvimento econômico e a alocação de recursos para atividades produtivas. Com isso, verifica-se uma forte correlação entre o desenvolvimento econômico e um mercado acionário ativo, pois este último proporciona pessoas, empresas, governos e outras entidades a investir suas poupanças em atividades produtivas.

O mercado de capitais brasileiros não é uma exceção ao elucidado acima, porém, apesar de crucial para economia brasileira, ele ainda não foi totalmente explorado. Com níveis de taxas de juros atingindo mínimas históricas até fevereiro de 2020, observa-se um fluxo de capital, principalmente de investidores brasileiros, para o mercado de ações nunca visto. Além dos níveis de SELIC, pode-se creditar este fluxo de capital ao desenvolvimento da infraestrutura e da diversificação da base de investidores pelo aumento do catálogo de instrumentos financeiros em circulação. Há ainda desafios como o pouco conhecimento popular sobre finanças, a baixa liquidez em instrumentos em mercados secundários e poucos estudos sobre as potencialidades da tecnologia neste mercado ainda com enorme potencial.

A seguir apresentaremos o conceito de análise técnica, bem como os índices utilizados nesse trabalho.

2.1.1 TEORIA DO MERCADO EFICIENTE

A teoria dos mercados eficientes tem sido um paradigma dominante ao descrever as variações dos ativos nos mercados especulativos. Working (1949) apresentou uma das primeiras versões da teoria e segundo o autor “Se é possível dada quaisquer combinações de circunstâncias para prever os preços futuros, quer dizer que as expectativas do mercado não estão precisas. Expectativas ideais por parte do mercado levariam em conta de fato toda a informação que permitiria a previsão correta das mudanças dos preços.”

Em outros trabalhos, o autor revisou sua definição de mercados futuros perfeitos para “...aquele em que o preço de mercado constitui a melhor estimativa possível a ser feita com a informação disponível no momento, de qual será o preço na data de vencimento/entrega do contrato futuro” (ver. Working 1962). Essa definição de um mercado futuro perfeito é na sua essência idêntica à definição de mercado eficiente dada por Fama (1970): “Um mercado cujos preço sempre refletem toda informação disponível é denominado “eficiente”. Desde este estudo de Fama foi publicado, essa definição de mercado serve como a definição padrão da literatura.

Uma definição mais prática de mercados eficientes foi dada por Jensen (1978) que afirma que “Um mercado é eficiente com respeito a informação em T_0 se é impossível ter lucro econômico com a mesma informação disponível em T_0 .” Dado que o lucro econômico é o retorno ajustado pelo risco deduzidos os custos transacionais, a definição de Jensen implica que a eficiência de mercado pode ser testada considerando os lucros líquidos e o risco de se negociar estratégias baseadas com a informação em T_0 .

Jensen (1978) agrupou as diversas versões de hipóteses de mercado eficiente em três formas:

- 1) Fraca: A informação disponível no momento é somente o histórico de preço do mercado em T_0 .
- 2) Semiforte: A informação disponível reflete toda a informação pública disponível no momento.
- 3) Forte: Toda a informação (pública ou privada) é conhecida no momento.

Portanto, a análise técnica, de certa maneira, testa a forma Fraca de eficiência de mercado devido ao uso de preços históricos. Testando a forma de eficiência empiricamente requer um modelo específico que pode descrever o processo de formação de preços quando estes representam toda a informação disponível. Neste cenário, o trabalho se propôs a utilizar-se de informações passadas e públicas para desenvolver um modelo capaz de prever o movimento dos preços de ações.

Ressalta-se o trabalho Lo (2005) onde cria a abordagem A Hipótese do Mercado Adaptativo em que modelos tradicionais de finanças moderna coexistem com modelos de finanças comportamentais. Implica que o grau de eficiência do mercado é relacionado aos fatores ambientais que caracterizam a ecologia de mercado como o número de competidores. Muitos estudiosos de finanças comportamentais citam como violações da racionalidade, que são inconsistentes com a Teoria do Mercado Eficiente, como: aversão a perdas, excesso de confiança, ancoragem e outros vieses são, na verdade, consistentes com um modelo em constante evolução de indivíduos se adaptando a mudanças do ambiente.

2.1.2 ANÁLISE TÉCNICA

A análise técnica é o processo de analisar tendências passadas de um investimento na tentativa de se determinar os prováveis comportamentos futuros. Ela envolve utilizar preços e dados passados para tomar decisões de investimento. Nos estudos de Schwager (1995) e Lo e Hasanhodzic (2010) é constatado que muitos *traders* e gestores entrevistados usam e confiam na análise técnica. Covel (2005) defende o uso exclusivo da análise técnica, citando como casos de sucesso grandes *hedge funds*. DeLong et al. (1990) apresentam teorias como a “*noise trader models*” que sugerem que operações utilizando análise técnica podem ser lucrativas sob um cenário de incertezas.

O estudo de Rotella (1992) disserta o fato que a análise técnica pode ser tanto usada de forma reativa quanto de maneira preditiva, ou seja, comportam-se de maneira passiva, após um acontecimento quanto se antecipando ao acontecimento. O preço de um ativo, em um determinado instante, representa um consenso pois simboliza o que um investidor concorda em negociar e depende primeiramente das suas expectativas. Se é esperado que o preço de um ativo irá subir, então ele será comprado, e se acredita que o preço do ativo caia, então deverá ser vendido. Tal questão causa o maior desafio na previsão do preço, uma vez que se referem às expectativas do ser humano, que são difíceis de serem quantificadas e previstas.

Conforme compilado por Hamilton (1922), a origem da análise técnica data do início do século XX, com a *Dow Theory*. Elaborada por Charles Dow, que nunca chegou a publicá-

la, consistia em acompanhar dois índices construídos pelo autor, a Média Industrial e a Média Ferroviária, compostos pelas principais *bluechips* disponíveis do mercado na época. A teoria consiste no fato de que o movimento destes dois índices determinaria as tendências do mercado como um todo. Portanto, há mais de 100 anos já se fazia análise de tendências com acontecimentos passados para previsões de preços de ações.

Tempos atrás, a discussão em torno do uso da análise técnica na precificação de ativos de renda variável foi bastante intensa no mundo acadêmico, como visto em Fama (1970). Deste ponto de vista, o seu uso foi encarado com certa desconfiança, sendo difícil imaginar o porquê de indicadores construídos a partir de comportamentos passados de um ativo específico ou mercado poderiam contribuir para investidores enxergarem retornos futuros de uma ação, por exemplo. Uma outra justificativa para o uso é que os padrões observados no mercado não refletem somente dados econômicos, mas fatores subjetivos, como a emoção e a lógica de cada agente.

Em Chew et al. (2003), conclui-se que dados extraídos a partir de indicadores técnicos podem desempenhar um papel importante na indicação do momento adequado de entrada e saída do mercado de ações. Aplicando estes indicadores, as empresas estudadas obtiveram resultados interessantes. Com isso, não é surpresa quando grandes *Hedge Funds*, empresas de gestão de ativos e bancos mantêm em seus quadros de funcionários, profissionais com conhecimento neste tipo de análise.

No trabalho de Menkhoff (2010), apresenta-se, apoiadas empiricamente, respostas para o porquê do uso da análise técnica. Na pesquisa foram entrevistados gestores e 87% viam a análise técnica como uma importante ferramenta para tomada de decisões. Outras conclusões do estudo foram: Em primeiro lugar que a análise técnica é importante e a maioria dos gestores de fundos de investimentos a utiliza como ferramenta de apoio aos negócios. Adicionalmente, a análise técnica é usada preferencialmente à análise fundamentalista para *trades* de curto prazo, os quais fazem previsões para poucos dias. Para previsões a partir de uma semana em diante, é mais usual o uso de análise fundamentalista, fato este percebido em todos os países estudados no trabalho. Terceiro, há evidências de informações heterogêneas no mercado financeiro, em contraste com o afirmado pela Teoria dos Mercados Eficientes. Analistas técnicos acreditam que há influências psicológicas e que elas são fatores importantes na precificação dos ativos no mercado financeiro. Acredita-se que existe um efeito manada (*herding effect*), quando agentes reagem aos comportamentos de outros agentes ignorando o comportamento do mercado como um todo e das análises fundamentalistas.

A popularização e o aumento da acessibilidade de técnicas de *machine learning* aumentou o número de trabalhos na literatura que versam sobre a junção destes dois temas. Várias ferramentas de data mining e inteligência artificial têm sido aplicadas a indicadores de análise técnica com objetivo de encontrar os melhores indicativos. Dash e Dash (2016) propuseram um modelo de *machine learning* e análise técnica para gerar sinais de compra ou venda na negociação de ações da bolsa de valores dos Estados Unidos. Choudhry e Kumkum Garg (2008) propuseram um sistema para prever preços de ações da bolsa de valores da Índia com *Support Vector Machine* utilizando-se de 35 indicadores técnicos e obtiveram bons resultados em termos de acurácia. Ratto et al. (2018) tentaram prever a direção do NASDAQ Stock Market, bolsa de valores americanas conhecida por negociar ações do setor de tecnologia. Além de indicadores de análise técnica os autores também utilizaram *Support Vector Machine* para efetuar suas previsões.

Os índices e indicadores de análise técnica que são descritos nos subitens a seguir servem como insumo de dados para o modelo de *machine learning* utilizado neste trabalho. Esses dados foram calculados para cada ativo utilizado no estudo em forma de Dados em Painel. O modelo então faz a análise dos indicadores e do preço histórico do ativo para prever seu valor futuro.

Irwin e Park (2007) constataam que 56 de 95 de estudos modernos sobre análise técnica produzem resultados positivos em relação ao uso da ferramenta, principalmente quando aplicados a mercados emergentes, como o Brasil.

Um enriquecimento do trabalho proposto por Irwin e Park (2007) é artigo de Nazário et al. (2017) em que 85 estudos foram revisados, classificados e agrupados em categorias e subcategorias. Os autores mostram quais estudos possuem maior influência, relevância acadêmica e aplicação na prática.

A seguir são listados os índices de análise técnica utilizados neste trabalho.

2.1.2.1 ÍNDICE DE FORÇA RELATIVA

Introduzido por Wilder Jr. (1978), o índice de força relativa é um indicador que mede a força das mudanças de preço com o objetivo de avaliar as condições de compra e venda de um determinado ativo. Conhecido como na sua sigla em inglês RSI (*Relative Strength Index*), o indicador oscila entre os valores de 0 a 100.

A interpretação comumente utilizada é quando o índice apresentar valores acima de 70, ele indica que o ativo está em níveis de sobrecompra e quando está abaixo de 30 indica que o ativo está sobrevendido.

O índice é calculado através da seguinte fórmula:

$$RSI = 100 - 100 / (1 + Média\ de\ Ganhos / Média\ de\ Perdas)$$

onde:

Média de Ganhos = Média dos ganhos dos períodos anteriores utilizados

Média de Perdas = Média das perdas dos períodos anteriores utilizados

Neste estudo foram utilizados três valores de RSI, calculados com valores de 6, 14 e 20 dias. O valor de 14 períodos foi proposto originalmente por Wilder Jr. (1978) e os valores de 6 e 20 dias são comumente utilizados no mercado.

2.1.2.2 BANDAS DE BOLLINGER

As Bandas de Bollinger (BB), introduzidas por Bollinger (2002), é um indicador de análise técnica que consiste em um conjunto de duas linhas desenhadas a uma distância de dois desvios-padrão da média-móvel do preço do ativo em questão. Um estreitamento no canal formado pelas bandas pode significar que, no momento, há um equilíbrio entre demanda e oferta e, após esse momento, haveria a possibilidade de uma forte tendência de alta ou baixa nos preços. Há indicações de tendência quando os preços ultrapassam os limites inferior e superior da banda. Caso o preço do ativo ultrapasse a banda superior, haverá uma tendência de alta do ativo. Entretanto, caso o preço fique abaixo da banda inferior, haverá então uma tendência de baixa. É calculado através das fórmulas:

$$\text{Banda Superior} = \text{MédiaMóvel}(PT, n) + m * \sigma[PT, n]$$

$$\text{Banda Inferior} = \text{MédiaMóvel}(PT, n) - m * \sigma[PT, n] ,$$

onde:

PT = Preço Típico que é a média simples entre preços da máxima e mínima além do preço de fechamento na data em questão.

n = Número de dias do período analisado

m = número de desvios-padrão utilizados

$\sigma[PT, n]$ = desvio padrão dos últimos n períodos dos Preços Típicos (PT)

Para este trabalho, $n = 20$ e $m = 2$, assim como no trabalho original de Bollinger.

2.1.2.3 MACD

O MACD (*Moving Average Convergence Divergence*) foi desenvolvido como uma ferramenta para poder identificar mudanças de tendências. O MACD é construído com base em médias móveis. Calcula-se através da subtração entre médias móveis ponderadas exponenciais, conforme demonstrado abaixo:

$$MACD(n) = EMA_k(i) - EMA_d(i)$$

onde: $k = 12$ e $d = 26$.

$$EMA_n(i) = \alpha * p(i) + (1 - \alpha) * EMA_n(i - 1)$$

$$\alpha = \frac{2}{1 + n}$$

onde: n é o número de dias e $p(i)$ é o preço do ativo em um dia i ; e EMA corresponde à média móvel exponencial e α ao seu fator de decaimento. Conforme utilizado em Murphy (1999), foram utilizados para os cálculos das médias móveis exponenciais de 12 e 26 dias. O sinal utilizado normalmente no MACD para indicar os momentos de entrada e saídas foi calculado conforme abaixo:

$$Sinal(n) = MMS_9(MACD(n))$$

onde: MMS é a média móvel simples de nove períodos passados do MACD.

Em Chong e Ng (2008) é indicado que o sinal de compra acontece quando o $MACD(n)$ é maior que zero e o valor do $MACD(n - 1)$ é menor ou igual a zero. Uma sinal de venda ocorre quando o sinal de $MACD(n)$ é menor que zero e o valor do $MACD(n - 1)$ é maior ou igual a zero.

2.1.2.4 MOMENTUM

É um indicador que mede a força da tendência de um índice ou do preço de um ativo. Proposto por Winder Jr. (1978), tem o objetivo de observar a variação diária do preço de um determinado ativo.

Definido como a diferença entre os preços de um determinado período e o preço de n períodos passados, conforme formulado abaixo:

$$\text{Momentum} = C - Cn$$

onde: C é último preço de fechamento, Cn é o último preço de fechamento de n períodos anteriores. Para este trabalho, $n = 2$, conforme o trabalho original de Winder Jr (1978).

Caso o resultado do cálculo do indicador demonstre um valor maior que zero, então indica um momento de compra, caso contrário, demonstra um valor menor que zero, então é um momento de venda.

2.1.2.5 OSCILADOR ESTOCÁSTICO

Apresentado na década de 1950 por Lane e testado por Huang et al (2015), o oscilador apresenta o estado do preço atual em relação a uma amplitude. A amplitude consiste na diferença entre os preços máximo e mínimo de um determinado período. Preços frequentemente próximos aos valores de amplitude máxima indicam uma força compradora e aqueles no fim da amplitude, força vendedora.

Lane explica que o estocástico varia entre 0 e 100 e, com isso, implementa áreas de sobrecompra e sobrevenda. O intervalo de valores entre 0 e 30 indica sobrevenda, enquanto a sobrecompra fica entre 70 e 100.

Neste oscilador são criados dois indicadores: O estocástico lento, conhecido como %K e o estocástico rápido, conhecido como %D.

São calculados através da seguinte formulação:

$$\text{Oscilador Estocástico Lento}(\%K) = C - Ll(14) / Hh - Ll(14)$$

$$\text{Oscilador Estocástico Rápido}(\%D) = SMA_3 \text{ do } \%K$$

onde: C é o preço de fechamento, $Ll(14)$ é o menor valor entre os preços mínimos dos últimos 14 períodos, $Hh(14)$ é o maior valor entre os preços máximos dos últimos 14 períodos e SMA_3 é a média móvel simples de 3 períodos, conforme proposto originalmente por Lane (1950).

2.1.2.6 CANAL DONCHIAN

Conforme Jain et al. (2007), o sistema Donchian foi criado em 1936 por Richard Donchian. Esta ferramenta é utilizada para identificar se a tendência do preço de um ativo irá continuar ou mudar.

Seu cálculo é simples e fácil de se efetuar:

Banda Superior: O maior valor para n períodos passados;

Banda Inferior: O menor valor para n períodos passados;

Banda Média: Média entre a banda superior e inferior.

Para este trabalho, $n = 20$, conforme originalmente proposto por Jain et al. (2007).

A indicação de compra aparece quando o preço de fechamento rompe a linha superior e a indicação de venda aparece quando o preço de fechamento rompe a linha inferior.

2.1.2.7 DESVIO PADRÃO, MÉDIA MÓVEL SIMPLES E MÉDIA MÓVEL EXPONENCIALMENTE PONDERADA

O desvio padrão é uma medida de dispersão e o seu valor reflete a variabilidade das observações em relação à média. O cálculo do desvio padrão é apresentado abaixo:

$$\sigma = \sqrt{\frac{1}{n} \sum_{t=1}^n (x_t - \mu)^2}$$

onde: σ é o desvio padrão, n é o número de elementos, x_t é o valor das variáveis aleatórias e μ é a média. Neste trabalho, os valores de x_t serão os preços de fechamento diários de cada período t .

A média móvel simples (MMS) pode ser calculada através do cálculo da média simples de um determinado período passado. A regra para este modelo é a seguinte: Para prever um dado na data t , no futuro, faz-se a média simples de n valores anteriores ao instante t .

O cálculo da média móvel simples é apresentado abaixo:

$$MMS = \frac{x_{t-1} + x_{t-2} \dots + x_{t-n}}{n}$$

onde: MMS é a média móvel simples, x é a variável a ser analisada, t é o período a ser analisado e n é a quantidade de total de períodos a serem observados. Foram utilizados os valores de 7, 15, 30 e 60 para n . Estes valores foram escolhidos por serem comumente usados no mercado.

Neste estudo foram utilizados tanto a média móvel simples dos preços de fechamento, quanto dos volumes de negociação diários de cada ativo.

Possui o mesmo objetivo da MMS, porém na média móvel exponencialmente ponderada (MMEP) utiliza-se de um fator de suavização no seu cálculo. Este fator tem o intuito de dar maior relevância para as informações mais recentes e menos para as mais antigas.

O cálculo da média móvel exponencialmente ponderada é apresentado abaixo:

$$MMEP_t = \lambda x_t + \lambda(1 - \lambda)x_{t-1} + \lambda(1 - \lambda)^2 x_{t-2} + \dots + \lambda(1 - \lambda)^m x_{t-m}$$

onde: $MMEP_t$ é a média móvel exponencialmente ponderada no instante t , λ é o fator de suavização, x é o valor a ser analisado, t é o período a ser analisado e m é a quantidade de total de períodos a serem observados. Para este trabalho, $m = 12, 26$ e 30 . Estes valores foram escolhidos por serem comumente usados no mercado.

2.1.2.8 RATE OF CHANGE

Conforme Chacko et al. (2015), a Taxa de Mudança, mais comumente referida como *Rate of Change* (ROC), é uma medida que mostra a variação entre um preço atual e outro preço passado. Em geral, os preços subirão enquanto o ROC for positivo e cairão caso o ROC for negativo.

Pode ser usado para se verificar divergências, como situações em que o ativo possa estar sobrecomprado ou sobrevendido. Os valores altos e baixos históricos do ROC funcionam como as regiões sobrecompradas e sobrevendidas, respectivamente. Abaixo segue a fórmula para seu cálculo:

$$ROC = \left(\frac{\text{Preço de Fechamento}_t - \text{Preço de Fechamento}_{t-n}}{\text{Preço de Fechamento}_{t-n}} \right) * 100$$

onde: t é o instante atual, n é número de períodos passados para observação e Preço de Fechamento é o preço de fechamento do ativo em determinada data. Para este trabalho, $n = 12$. Este valor foi escolhido por ser comumente utilizado no mercado.

2.1.2.9 PIVOT POINTS

Pivot point, conforme Satchwell (2010), é uma ferramenta de análise técnica usada para determinar a tendência geral do mercado em diferentes janelas temporais. Seu cálculo é simples e se traduz em uma média simples entre os preços de abertura, fechamento e médios de um determinado dia. Caso, no próximo dia, o preço de negociação esteja acima do valor encontrado para o *Pivot Point*, há indicação de um mercado com tendência de alta (*bullish*) e, caso esteja abaixo, há indicação de que o mercado está com tendência de baixa (*bearish*). É calculado da seguinte forma:

$$Pivot\ Point = \frac{Preço\ de\ Abertura + Preço\ de\ Fechamento + Preço\ Médio}{3}$$

2.1.2.10 MASS INDEX

Apresentado por Dersey (1992), o *Mass Index* foi desenvolvido para identificar mudanças de tendência, medindo as aberturas e fechamentos entre as diferenças das máximas e mínimas dos preços. Quando a diferença aumenta (diminui), o *Mass Index* aumenta (diminui). Uma tendência muda quando o índice depois ter atingido o valor 27, diminui e atinge patamares inferiores a 26,5. O índice é obtido através do cálculo da média móvel exponencial dos últimos 9 dias dividida pela média móvel exponencial de 9 dias da média móvel exponencial dos últimos 9 dias e, por último, somando-se esta razão pelos últimos 25 dias. Calcula-se da seguinte forma:

$$\text{Razão da MME} = \frac{\text{MME de 9 dias}(\text{Máxima} - \text{Mínima})}{\text{MME de 9 dias da MME de 9 dias}(\text{Máxima} - \text{Mínima})}$$

$$\text{Mass Index} = \text{Razão da MME}_t + \text{Razão da MME}_{t-1} + \text{Razão da MME}_{t-2} + \dots + \text{Razão da MME}_{t-n}$$

onde: Razão da MME é a razão entre a *MME de 9 dias (Máxima – Mínima)* é a média móvel exponencial de 9 períodos da diferença entre os preços máximo e mínimo de cada período e *MME de 9 dias da MME de 9 dias (Máxima – Mínima)* é a média móvel exponencial de 9 períodos da média móvel exponencial de 9 períodos da diferença entre os preços máximo e n é número de períodos passados para observação. Estes valores foram escolhidos por serem comumente usados no mercado.

2.1.2.11 TRUE STRENGTH INDEX

Apresentado por Blau (1991), o *True Strenght Index* (TSI) usa diferenças para medir as mudanças de preço. O TSI utiliza uma função de diferenciação para medir o momentum e uma função de média para correlacionar o momento à tendência do preço. Quando o valor do índice é positivo, indica a alta do ativo e quando negativo, indica a baixa no preço do ativo. É calculado através da seguinte fórmula:

$$\text{TSI}(\text{preço}, n, m) = \frac{\text{MME}(\text{MME}(\text{momentum}, n), m)}{\text{MME}(\text{MME}(|\text{momentum}|, n), m)}$$

onde: $\text{MME}(\text{momentum}, n)$ é a média móvel exponencial do momentum por n períodos, e $\text{MME}(\text{MME}(|\text{momentum}|, n), m)$ é a média móvel exponencial da média móvel exponencial por m períodos. Para este trabalho, $n = 25$ e $m = 13$. Estes valores foram escolhidos por serem comumente usados no mercado.

2.1.2.12 LINHA DE ACUMULAÇÃO/DISTRIBUIÇÃO

Segundo Thomsett (2010), o método, conhecido popularmente com A/D, foi criado por Marc Chaikin e desenvolvido com o objetivo de determinar, dada uma certa tendência, qual força, compradora ou vendedora, é a mais dominante e identificar o quanto forte e rápida essa tendência evolui. É calculada através da seguinte fórmula:

$$Money\ Flow\ Volume_t = \frac{(Fechamento - Mínima) - (Máxima - Fechamento)}{(Máxima - Mínima)} * Volume$$

$$A/D_t = A/D_{t-1} + Money\ Flow\ Volume_t$$

onde: $Fechamento_t$ é o preço de fechamento de um determinado ativo na data t , $Mínima_t$ é o preço mínimo atingido pelo ativo na data t , $Máxima_t$ é o preço máximo de um ativo atingido na data t , $Money\ Flow\ Volume_t$ é fluxo financeiro transacionado na data t , A/D_{t-1} é linha de acumulação/distribuição do período imediatamente anterior a t .

A tendência de alta é confirmada por constantes valores crescentes do indicador e a tendência de baixa é confirmada por constantes valores decrescentes do indicador. A indicação da fraqueza da tendência de alta ocorre quando os preços sobem para novos máximos e não são confirmados pela dinâmica do indicador. Quando isso acontece é possível uma reversão dos preços para baixo. A indicação da fraqueza da tendência de baixa ocorre quando os preços caem para novos mínimos e não são confirmados pela dinâmica do indicador. Quando isso ocorre, é possível uma reversão dos preços para cima.

2.1.2.13 OSCILADOR DE CHAIKIN

Também criado por Marc Chaikin, é um indicador de volume baseado na linha de acumulação/distribuição. Quando o indicador é positivo, indica uma forte força compradora e quando negativo, indica uma forte força vendedora. O valor do oscilador é derivado da subtração das médias móveis exponenciais longa e curta da linha de acumulação/distribuição (conforme o item 2.1.1.12). Pode ser usado como uma maneira de se medir a taxa de aceleração dessa linha.

É calculado através da seguinte fórmula:

$$Oscilador\ de\ Chaikin = MME\ de\ 3\ dias\ de\ A/D - MME\ de\ 10\ dias\ de\ A/D$$

onde: *MME de 3 dias de A/D* corresponde à média móvel exponencial de 3 dias da linha de acumulação/distribuição (conforme o item 2.1.1.12) e *MME de 10 dias de A/D* corresponde à

média móvel exponencial de 3 dias da linha de acumulação/distribuição. Estes valores foram escolhidos por serem comumente usados no mercado.

2.1.2.14 ÍNDICE DE FORÇA

Mais conhecido pela sua forma em inglês, *Force Index*, foi introduzido por Elder (1993), e mede a força compradora ou vendedora por trás de cada tendência. Quando os valores do índice forem negativos, indicam força vendedora com desvalorização do preço do ativo. Alternativamente, quando o índice for positivo, ele indica força comprado com valorização do preço do ativo. Esse índice demonstra a direção das mudanças e a extensão das mudanças de preço e o volume negociado, todos esses combinados.

É calculado através da seguinte fórmula:

$$Force\ Index = Volume_t * (Fechamento_t - Fechamento_{t-1})$$

onde: $Volume_t$ corresponde ao volume do ativo negociado em uma data t , $Fechamento_t$ corresponde ao preço de fechamento do ativo em uma data t .

2.1.2.15 EASE OF MOVEMENT

Desenvolvido por Arms (1996), é um indicador que mostra a relação entre as mudanças no preço de um ativo e o seu volume de negociação, verificando qual a maior probabilidade para o próximo movimento no preço, subir ou cair. Quando o indicador permanecer por um período estabilizado com valores positivo, indica uma alta nos preços. Alternativamente, quando o indicador permanecer estabilizado com valores negativos, indica uma baixa nos preços. Primeiramente, calcula-se duas variáveis para a obtenção da métricas. A primeira, chamada de *box ratio* (BR) e a segunda, chamada de *mipoint move* (MPV). Ambas informações são essenciais no cálculo do *Ease of Movement*, conforme demonstrado:

$$BR_t = \frac{Volume_t}{(Preço\ Máximo_t - Preço\ Mínimo_t)}$$

$$MPV_t = \left(\frac{Preço\ Máximo_t - Preço\ Mínimo_t}{2} \right) - \left(\frac{Preço\ Máximo_{t-1} - Preço\ Mínimo_{t-1}}{2} \right)$$

$$Ease\ of\ Movement_t = \frac{BR_t}{MPV_t}$$

onde: $Volume_t$ é o volume negociado total na data t , $Preço\ Máximo_t$ é o maior valor observado para um ativo na data t , $Preço\ Mínimo_t$ é o menor valor observado para um ativo na data t .

2.2 MACHINE LEARNING

De acordo com Shalev-Shwartz e Ben-David (2014), o termo *Machine Learning* pode ser entendido como uma técnica de análise de dados baseada na habilidade de uma máquina aprender de experiências passadas, e melhorar conforme a sua performance sem intervenções antrópicas. Esse sistema atua com a detecção automática de padrões relevantes em dados. Nas últimas duas décadas o uso dessa ferramenta se tornou mais comum em qualquer atividade que envolve extração de informação de uma grande série de dados. As ferramentas de *machine learning* têm o objetivo de prover soluções com a habilidade de “aprender” e “adaptar”. Podemos então entender *Machine Learning* como uma ramificação, ou junção da estatística e computação, reunindo métodos que tem como objetivo automatizar o processo de aprendizado de uma certa base de dados.

Finnie et al. (2010) revisam a literatura a respeito de técnicas de *machine learning* aplicadas a previsão de séries temporais de finanças. Os autores mostram que 67% dos trabalhos analisados efetuam previsões com base em dados com intervalos de tempo diários, 75% dos trabalhos algum tipo de dados passados em suas previsões: preços de fechamento, abertura, volume e indicadores de análise técnica. 15% dos trabalhos analisados utilizaram a técnica *buy and hold* como técnica de alocação dos portfólios obtidos com redes neurais. Cerca de 80% dos portfólios obtidos nos estudos tiveram uma performance superior ao *benchmark* proposto. Todas as técnicas citadas foram utilizadas neste trabalho.

Gu et al. (2018), estudaram a existência ou não de previsibilidade no mercado de ações dos Estados Unidos, utilizando algoritmos de *machine learning*, e chegaram à conclusão de que, para os modelos de árvores de decisão e redes neurais, há evidências de que existe previsibilidade nas alterações dos preços das ações.

Kim (2003) utiliza-se do SVM (*Support Vector Machine*), uma técnica de *machine learning*, para prever o comportamento do KOSPI (*Korea Composite Stock Price Index*), índice de ações da Coreia do Sul, utilizando-se também de indicadores de análise técnica. Katon, McNally e Roche (2018) estudaram o desempenho do algoritmo utilizado neste trabalho, LSTM (*Long-Short Term Memory*), em paralelo com um modelo popular de previsão de séries temporais, ARIMA (*Autoregressive integrated moving average*) para previsão do preço do *Bitcoin*. Como esperado, o modelo de aprendizado não-linear LSTM obteve um desempenho superior ao ARIMA.

Um modelo semelhante ao proposto neste trabalho foi construído por Obeidat et al. (2018), ao construir uma carteira composta de ETFs (*Exchange-traded funds*) utilizando

previsões de retorno destes fundos oriundas da rede LSTM e sua alocação baseada na técnica de MVO. Posteriormente compara este modelo a outras técnicas de alocação (CAPM⁴, 3FM⁵ e EQWT⁶). O portfólio obtido através do LSTM possui um melhor desempenho comparado ao construído com outras técnicas. Já Nelson et al. (2017) também estudaram a aplicabilidade do algoritmo LSTM com preços históricos de ações negociadas na bolsa de valores brasileira (B3). O objetivo era saber se o preço da ação iria subir ou descer nos próximos 15 minutos. O resultado da pesquisa mostra que o algoritmo se provou eficaz em prever retornos de ativos de renda variável. Cui et al. (2018) estudaram o movimento do índice da bolsa ações da China (CSI 300), através da previsão do seu comportamento ao longo do ano de 2016 com um modelo de redes neurais multifiltros (MFNN) e, mais uma vez, o modelo de *Machine Learning* mostrou-se eficaz na proposta de previsão de séries temporais.

Várias técnicas de *Machine Learning* podem ser utilizadas para previsões de variáveis a partir de series temporais. Assim como Varian (2014), fizemos o gráfico do Google Trends para a expressão “*machine learning time series*” (Gráfico 1). No gráfico, verifica-se um crescente interesse pelo tema nos últimos cinco anos.

⁴ CAPM: *Capital Asset Pricing Model*. Apresentado por Sharpe (1970) descreve a relação entre o risco sistemático o retorno esperado em ativos, principalmente ações.

⁵ 3FM: *Fama and French Three Factor Model*: Desenvolvido por Fama e French (1992), objetiva descrever o comportamento dos retornos de ações através de três fatores: risco de mercado, a maior performance de ações *small-caps* em relação a ações *large-caps*, e a maior performance de ações com altos valores de *book-to-market* contra ações com baixos valores de *book-to-market*.

⁶ EQWT: *equally weighted portfolio rebalancing* é uma ponderação que dá o mesmo peso para todos os ativos de um portfólio.

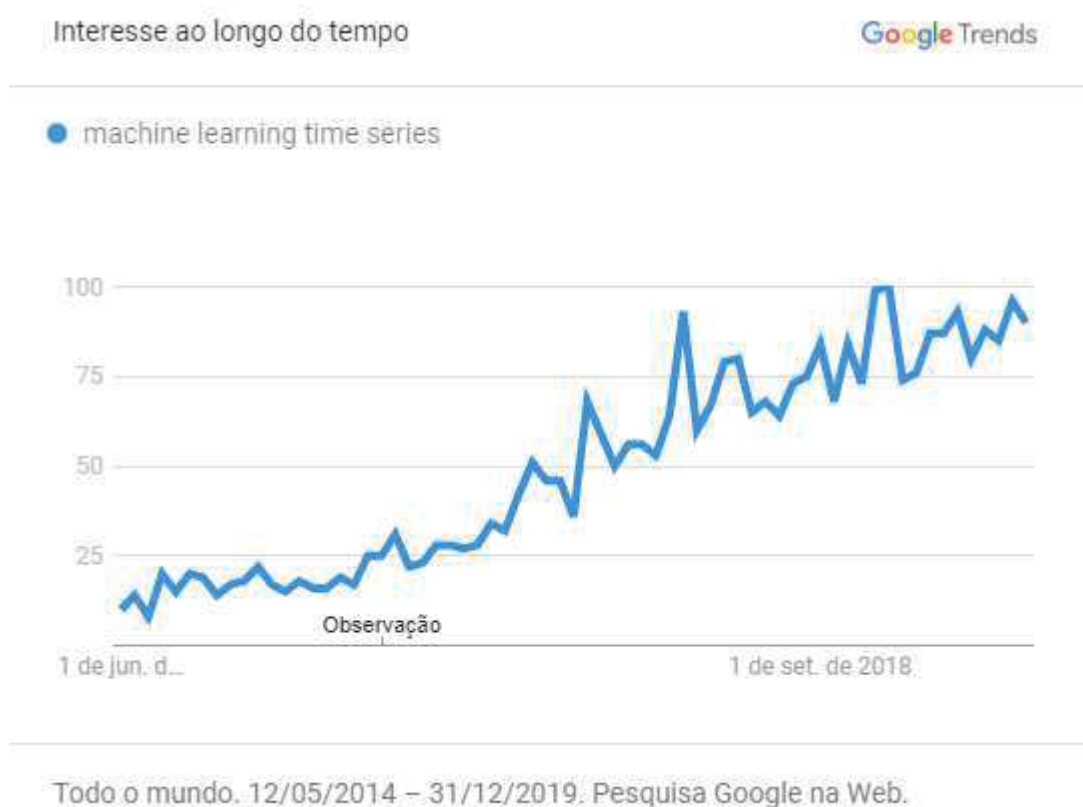


Gráfico 1 - Gráfico Google Trends: Busca pelo termo “machine learning time series” nos últimos 5 anos. Período: 12/05/2014 a 31/12/2019

Kleinberg (2015) argumenta que também há uso em campos onde o resultado do modelo de *Machine Learning* contribui para tomada de grandes decisões. Ele fornece, como exemplo, o campo da medicina, onde um modelo sugere se um paciente idoso deve ou não efetuar uma cirurgia no quadril, baseado em dados de outros pacientes em situações semelhantes, com características diversas.

A análise preditiva tem como objetivo principal estimar algo no futuro que hoje é incerto, a fim de guiar a tomada de decisões hoje. Desta forma, *Machine Learning* tem o propósito de descobrir padrões nos dados para balizar a tomada de decisões, com previsões geradas que comumente seriam difíceis de o cérebro humano realizar, devido à complexidade dos dados.

De acordo com Shalev-Shwartz (2014), o termo *Machine Learning* refere-se à detecção de significado em dados. Devido à complexidade dos padrões a serem detectados, técnicas tradicionais ou o próprio olho humano não são capazes de detectar.

Conforme Frank et al. (2011), *Machine Learning* é utilizada para extrair informações cruas de uma série de dados em uma forma compreensiva que pode ser usada para diversos propósitos.

Conforme Lorena e Carvalho (2007), as técnicas de *Machine Learning* utilizam um princípio denominado indução, onde é permitido alcançar resultados abrangentes a partir de um certo conjunto de dados. De acordo com Jung et al. (2018), os métodos de ML podem ser categorizados em três tipos:

- (i) aprendizado supervisionado;
- (ii) aprendizado não supervisionado; e
- (iii) aprendizado por reforço (*reinforcement learning*), que busca iterativamente uma localização ideal das variáveis de entrada que geram a maior recompensa, ou seja, otimiza uma determinada função de "recompensa" não usando nenhum conjunto de treinamento (por exemplo, modelos de programação dinâmica, sarsa, Q-learning).

O método não supervisionado envolve encontrar grupos de observação (*clusters*) que são similares em relação a seus regressores. É comumente utilizado para processamento de dados de vídeo, imagem ou texto. O produto de um modelo não supervisionado é uma partição dos dados de entrada, onde cada partição funciona como um grupo, identificando-se similaridades de acordo com algum critério. Portanto, a sua utilização não será escopo deste trabalho.

Métodos supervisionados, conforme Pellucci et al. (2011), envolvem tipicamente o uso de vários mecanismos e regressores para prever uma variável dependente. Esta tarefa é realizada através do fornecimento de uma referência do objetivo a ser alcançado, ou seja, um conjunto de dados para aprendizagem do algoritmo. Este conjunto de dados, conhecido como dados de treinamento, são exemplos com diversas entradas e com uma saída esperada. O algoritmo, através da elaboração de uma função que mais aproxima os dados de entrada com os dados de saída, objetiva que a representação obtida seja capaz de produzir novas saídas a partir de entradas que não foram apresentadas anteriormente.

Há uma grande variedade de métodos de *Machine Learning*, como: regressão linear, regressão regularizada (LASSO, Ridge e Elastic Net), *Random Forest*, Árvores de Regressão, *Support Vector Machines* (SVM), Redes Neurais, entre outros.

2.2.1 REDES NEURAIS

As redes neurais podem ser entendidas como um método de regressão, ou de modelagem, em dois estágios. No entanto, ao contrário dos demais métodos (vide item 2.2), as redes neurais estimam relações não lineares nas suas etapas intermediárias.

Redes Neurais são elaboradas com a intenção de imitar a maneira que um cérebro humano funciona e são o pelotão de frente do crescimento exponencial no campo da Inteligência Artificial. Foram discutidas em trabalhos mais genéricos e artigos como Chen, Frias-Martinez, Macredie e Magoulas (2006), Filstad, Gottschalk, Glomseth e Solli-Sæther (2011) e Yaqoob et al (2016).

Pensamentos humanos são sempre encadeamentos de outros de pensamentos, palavras, eventos e experiências, ou seja, há uma continuidade. Redes Neurais Recorrentes também funcionam desta maneira. São redes com *loops*, permitindo que a informação que trafega através dela persista.

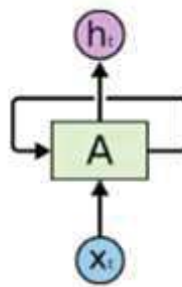


Figura 1 - Estrutura simplificada de uma Rede Neural Recorrente.

Na Figura 1, há uma versão simples de como funciona uma rede neural. O ponto A recebe um input de x_t e tem como *output* h_t . O *loop* na imagem ilustra a informação a ser passada de uma etapa para próxima da rede. Uma rede neural recorrente pode ser pensada como o conjunto de múltiplas cópias da mesma rede, cada uma passando uma mensagem ao seu sucessor.

2.2.2 LONG-SHORT TERM MEMORY (LSTM)

Long Short Term Memory (LSTM), introduzido por Hochreiter e Schmidhuber (1997), são redes neurais recorrentes modificadas. Essas redes possuem a propriedade de lembrança seletiva de padrões por longos períodos de tempo, sendo capaz de realizar processos comuns a redes neurais humanas e são utilizadas em trabalhos de classificação, processamento e previsão de séries temporais. É uma arquitetura de rede neural recorrente que tem a habilidade de armazenar valores em intervalos de tempo arbitrários, com isso, detém a capacidade de guardar informações em dados por períodos longos quando se trabalha com séries temporais.

Uma camada da LSTM consiste em um conjunto de blocos interligados, conhecidos como células de memória. Estes blocos podem ser pensados como diferentes versões de um *chip* de memória de um computador. Cada um contém uma ou mais células de memória conectadas a três unidades multiplicativas, o *input gate*, *output gate* e o *forget gate*. Estas três unidades têm a capacidade de efetuar a escrita, leitura e reinicialização de dados contendo informações para as células.

Esta corrente de dados é feita mantendo-se o deslocamento do erro constante através das unidades especiais, os *gates* de uma célula, que permitem ajustes de pesos de informações recentes ou antigas, ou seja, o aumento da relevância do mesmo para a série, da mesma forma que a diminuição da relevância da sequência quando a informação não é necessária, de acordo com o modelo, representando uma abstração da sequência de dados.

Efetuando uma analogia com o entendimento de um ser humano pode ser dito que assim como lemos um texto, compreendemos o sentido de cada palavra de acordo com o sentido da palavra anterior. O entendimento completo não se baseia apenas no conhecimento de cada palavra individualmente e sim na fluidez do texto como um todo. Nesse mesmo sistema, funcionam as previsões de uma rede LSTM, uma vez que seus entendimentos estão condicionados à experiência passada dos *inputs* recebidos.

As redes LSTM possuem a capacidade de armazenar informações por longos períodos ao processar uma sequência temporal. Sendo esta característica extremamente relevante quando se trata do estudo realizado neste trabalho com sequência de dados de mais de 10 anos. O funcionamento da célula de uma rede LSTM é descrito e ilustrado na Figura 2 e é representado pelas equações abaixo, conforme Santos (2019):

$$\begin{aligned}
f_t &= \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \\
i_t &= \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \\
\tilde{C}_t &= \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \\
C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\
o_t &= \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \\
h_t &= o_t * \tanh(C_t)
\end{aligned}$$

onde: x_t é o *input* da rede na data t , h_t é o *output* da célula na data t , σ denota a função de ativação logística, C_t denota o estado da célula no momento t , \tilde{C}_t representa o candidato de estado para na data t . Adicionalmente há 3 *gates* nas células LSTM que são: *forget gate* f_t , *input gate* i_t e o *output gate* o_t . As constantes w_d , w_i , w_o e w_c são os pesos dos *gates* de esquecimento, entrada, saída e da própria célula, respectivamente, ou seja, o quanto de relevância é dada a um dado. As constantes b_f , b_i , b_o e b_c representam as extremidades do *forget*, *input* e *output gates* e da própria célula, respectivamente, conforme Greff et al. (2016).

Conforme Badola et al. (2018) a célula da rede LSTM é descrita pela figura 2.

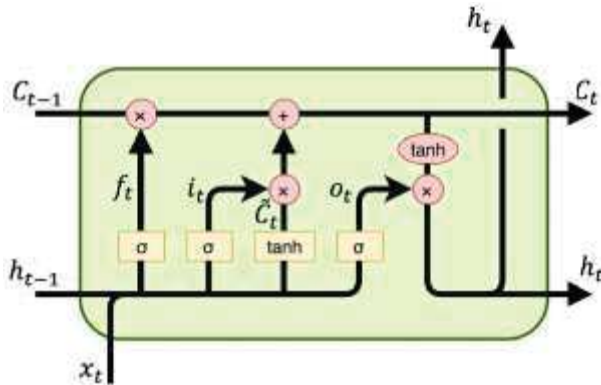


Figura 2 - Representação de uma célula LSTM.

Fonte: Ganesh (2019)

O *input gate* (i_t) regula o quanto será memorizado no interior da célula, na data t . Já o *output gate* (o_t) controla o quanto de informação será removida da célula, na data t . Greff et al. (2016) analisam a mesma estrutura da rede LSTM apresentada neste trabalho e verificam que é possível obter-se bons resultados quando se aplica esse modelo à previsões de séries temporais. Conforme o esquema apresentado, percebe-se que neste formato de células que as redes LSTM conseguem aprender relações de longo prazo a partir de uma estrutura de dados temporal.

2.2.3 PREVISÕES

Apesar de não haver referência direta na literatura sobre a janela de previsão indicada para cada modelo, cada negócio ou cada área, foi possível através de análises de trabalhos e pesquisas, encontrar janelas em que o modelo responde melhor.

No trabalho de Majidi et al. (2018) testa-se diversas janelas de previsão para preços de ações em diversos modelos de *machine learning*. Constata-se que janelas de previsão acima de dez passos (*steps*) resultam em modelos com erros elevados.

Analisando os trabalhos de Holder *et al.* (2018), Finkelstein e Jeong (2017), Cheng et al. (2006), Dong e Li (2017) e Majidi et al. (2018), que utilizam-se de modelos de *machine learning* para previsões diversas, em vários campos de atuação diferentes, verifica-se que todos os autores entram em um consenso quanto ao uso de janelas de previsões de até dez períodos

Um estudo feito por Fränkel (2018) analisa diferentes janelas de previsão para resolução de problema de regressão de séries temporais. O autor conclui que os resultados para previsões de 1 e 2 períodos, apesar de apresentarem um erro quadrático médio baixo em relação a previsões que utilizam 3 até 96 janelas de previsão, possuíam alta autocorrelação entre as previsões e a base de dados. O autor indaga na conclusão da eficácia de modelos de previsão com janelas para 1 período. Yunpeng et al. (2017) também fazem o estudo de janelas de previsão para a rede LSTM em problemas de regressão de séries temporais. Os autores também apontam o problema de elevada autocorrelação entre os as previsões e os dados informados ao modelo para previsões de 1 período a frente.

As janelas de previsão escolhidas foram de 3, 4 e 5 dias. Constata-se pela revisão da literatura mencionada que as janelas utilizadas foram as mais adequadas na utilização deste tipo de problema de séries temporais pois através destas os modelos finais possuíam o menor erro quadrático médio, quando comparados a outras janelas de previsão. O erro quadrático médio (EQM) é definido abaixo:

$$EQM = \frac{1}{n} \sum_{n=1}^n (Y_i - \hat{Y}_i)^2$$

onde: Y_i é o preço real observado de uma ação i , \hat{Y}_i é o preço previsto de uma ação i e n é a quantidade de períodos.

As previsões, dentro do modelo de *Machine Learning* funcionam da seguinte maneira:

Primeiro, o modelo é treinado com a base de dados de treinamento, que consistem nos retornos, volumes e variáveis de análise técnica passado, no período entre janeiro de 2008 a dezembro de 2018. Todavia, apesar de não ser definido como regra empírica, para adequar os dados à modelagem é recomendado, conforme James et al. (2013), que todas as variáveis da base de dados sejam normalizadas.⁷ A normalização consiste no seguinte processo:

- Normalização:

$$z_i = \frac{x_t - \mu}{\sigma}$$

⁷ Normalização, no contexto deste estudo, é aplicada para que variáveis de grandezas distintas, como volume e preço, que podem possuir grandes diferenças em seus valores, que para um modelo multivariável como este pode influenciar o resultado dado a sutuosidade destas diferenças.

A normalização tem o objetivo de manter as bases de dados organizadas e assim reduzir a redundância de dados, aumentar a integridade e o desempenho do modelo.

- Média:

$$\mu = \frac{1}{n} \sum_{t=1}^n x_t$$

- Desvio Padrão:

$$\sigma = \sqrt{\frac{1}{n} \sum_{t=1}^n (x_t - \mu)^2}$$

onde: x_t é o valor não normalizado do atributo x em uma data t , z_t é o valor normalizado do atributo x em uma data t , μ é a média aritmética do atributo x , σ é o desvio padrão do atributo x e n é o total de datas da base de dados no período, neste trabalho foi utilizado um valor de 2711 para n

Nesta etapa de treinamento, assim como os dados de *input*, os parâmetros do modelo são definidos. No total são construídos 3 modelos de previsão, cada um observando os mesmos dados de treinamento, porém distintos entre si na variável da janela de previsão (neste trabalho 3, 4 e 5 dias). Esta mudança, naturalmente, provoca a criação de três modelos que produzirão previsões de retornos diferentes.

Após os três modelos serem construídos, as previsões começam a ser efetuadas. São fornecidos os dados de teste, novos aos modelos construídos. Estes dados consistem nos dados passados do período de janeiro de 2019 a setembro 2019. Todas as variáveis, assim como na fase de treinamento, são fornecidas, com exceção do valor dos retornos das ações no ano de 2019, pois estes serão os valores a serem previstos, pelos três modelos. Portanto, o modelo recebe os valores de x_t (volume de negociação diária e dados de análise) e prevê y (retornos das ações).

A acurácia do modelo LSTM é medida com base nos valores previstos através do fornecimento de valores passados. Isto ocorre alimentando o modelo com a última previsão que foi alimentada o modelo LSTM, que será a última entrada de dados da base de treinamento, nós então faremos uma previsão que representará o próximo preço de fechamento ajustado. Guarda-se os valores previstos que serão chamados P0 e começa-se a prever o próximo valor de preço de fechamento ajustado alimentando P0 ao modelo LSTM e teremos o valor previsto P1. Este processo iterativo continuará passo a passo até que o número de previsões desejado seja alcançado. Evidentemente, este processo prossegue após concluído primeira janela de previsão, observando os valores que já foram previstos, por toda a extensão da base de dados de teste.

2.3 CONSTRUÇÃO DE PORTFÓLIO

Modelos de otimização de portfólios, utilizando-se de técnicas de aprendizado de máquinas são bastante utilizados em Finanças. Muitos autores utilizaram redes neurais para previsão de séries temporais de finanças em modelos de otimização de portfólio, como: Obeidat et al. (2018), Vukovic et al. (2020) e Vyklyuk et al. (2013).

Em “*Portfolio Selection*”, trabalho elaborado em 1952 por Markowitz, originou-se a análise de média-variância de carteiras, revolucionando a maneira em que o processo de alocação de ativos era feito. Passando para um processo novo, denominado otimização, Markowitz elaborou a ideia de que investidores deveriam considerar a relação entre retorno esperado e risco incorrido para um determinado investimento. Com isso, poderiam avaliar qual a melhor alocação de carteira. Entende-se “melhor” como para uma determinada variância, escolher a carteira com o maior retorno esperado possível. Pode-se pensar também através de outra ótica: para um determinado retorno esperado, escolher a carteira com a menor variância possível. Estas ideias foram revolucionárias do ponto de vista da mudança de foco na alocação de investimentos, onde era observado por uma visão sistemática. Tendo assim uma visão voltada para diversificação e não observando somente ativos individualmente, dando mais enfoque à relação entre o risco e o retorno. Pela ótica do risco, a covariância entre ativos e não mais individuais começou a ganhar mais relevância.

Apesar do seu impacto positivo, ainda há desconfiança do modelo por parte de gestores. Esta desconfiança se dá na dificuldade na estimação precisa dos retornos esperados e na obtenção da matriz de covariância para estes ativos. Modelos tradicionais utilizam, geralmente, máxima verossimilhança para previsão de retorno, conforme DeMiguel e Nogalas (2009). Estes modelos, apesar de eficientes, mostram-se sensíveis à desvios de distribuição, pois comumente assumem a hipótese de uma distribuição normal.

Almeida et al. (2001) elaboraram um modelo com base somente em preços históricos de 66 ações que compunham o Ibovespa. Foram comparados dois modelos: o primeiro de alocação clássica de MVO, onde só se observava variações de preço passadas, e o modelo de rede neural Autorregressivo, que efetuou previsões para os preços dos ativos em questão. O resultado da carteira que utilizou redes neurais foi superior ao modelo tradicional de Markowitz, em termos de rentabilidade.

Corroborando para a utilização de *Machine Learning* para seleção de carteiras, há diversas evidências na literatura de que a suposição de normalidade dos retornos nem sempre é eficaz. Estas presunções podem levar a erros de estimação que impactam diretamente na

construção dos portfólios. Nos estudos realizados por Michaud (1989), Best e Grauer (1991) e Mendes e Leal (2005), constata-se que sutis mudanças nos *inputs* para os cálculos matriciais da covariância e os cálculos dos retornos esperados impactam significativamente as distribuições dos pesos entre os ativos na carteira ótima. Revisitando o trabalho efetuado por DeMiguel e Nogales (2009), verifica-se que estas carteiras, construídas com estimadores enviesados, são inferiores em desempenho a carteiras cujos estimadores não apresentam viés.

2.3.1 ÍNDICE DE SHARPE

Com o mesmo nome do seu criador, o índice de Sharpe, introduzido por Sharpe (1966), é a medida de retorno ajustado pelo risco usada comumente para comparar ao desempenho de investimentos.

O índice de Sharpe de um portfólio também pode ser entendido como a relação entre o retorno e o risco, sendo este ligado à medição do desvio padrão. O quadrado do desvio padrão é a variância, medida utilizada por Markowitz. O índice de Sharpe mede o retorno esperado pelo portfólio acima da taxa livre de risco por unidade de risco total assumido, ou seja, por unidade de volatilidade dos retornos do portfólio:

$$S_p = \frac{E(R_p) - R_f}{\sigma(R_p)}$$

onde: S_p é o índice de Sharpe do portfólio, $E(R_p)$ é o retorno esperado do portfólio, R_f é o retorno do ativo livre de risco e $\sigma(R_p)$ é o desvio padrão dos retornos do portfólio.

Então, por que Sharpe utilizou o desvio padrão para ajustar o excesso de retorno de um investimento? Por exemplo, uma ação X gera um retorno de 20% e a ação Y gera um retorno de 18%. Analisando somente estes dados, presume-se que a ação X é mais interessante de se investir que a ação Y . Dado uma taxa livre de risco de 10%, o desvio padrão da ação X em 10% e o desvio padrão de Y em 5%. O índice de Sharpe da ação X é 1, enquanto o índice de Sharpe da ação Y é 1,6. Dados estes cálculos, conclui-se que a ação Y possui um melhor retorno ajustado ao risco.

O modelo proposto neste trabalho traz a união de conceitos de ambas as teorias apresentadas. A diversificação de Markowitz para otimização de portfólio e a utilização da métrica do índice de Sharpe, para obter-se o melhor resultado de retorno possível ajustado ao risco.

2.3.2 MODELO DE OTIMIZAÇÃO MÉDIA-VARIÂNCIA

A teoria moderna de portfólio foi desenvolvida por Markowitz em 1952. O estudo estabeleceu como principal vertente a relação risco-retorno dos investimentos.

A teoria afirma que maiores (menores) riscos estão associados com as possibilidades de se obterem maiores (menores) retornos. Com isso, dados dois portfólios, investidores avessos ao risco irão escolher aqueles que, dados os mesmos retornos, possuem o menor risco. Portanto, um investidor racional só irá assumir mais risco em seus investimentos se for recompensado com mais retorno.

Outro aspecto trazido no estudo foi o conceito de diversificação. A teoria moderna de portfólio diz que se um investidor investe em mais de uma ação, ele se beneficia do fator diversificação, reduzindo o risco de sua carteira.

Tratando do retorno de uma carteira podemos afirmar que seu retorno será a média ponderada dos ativos que a compõem, mas quando se trata de risco não é possível tratar com a mesma simplicidade. O risco deve contemplar a correlação dos movimentos dos ativos dentro da carteira. A diversificação afirma que se pode obter menos risco do que quando se investe em um único ativo, com o mesmo ou maior retorno, otimizando sua alocação.

Conforme já explicado, a otimização pelo método média-variância (MVO - *Mean Variance Optimization*, em inglês) maximiza o retorno a partir de um certo risco ou minimiza o risco para um certo retorno. O MVO necessita das previsões dos retornos esperados dos ativos, seus desvios-padrão e de sua matriz de covariância para encontrar as alocações e, finalmente, o portfólio eficiente.

O método de MVO foi usado neste trabalho para gerar a alocação ideal. Não será considerada a venda a descoberto e nem o aluguel ou empréstimo de ativos sem risco. Em ambos modelos a matriz de covariância é a medida utilizada para calcular o risco da carteira.

Os desvios-padrão e a matriz de covariância, neste trabalho, são calculados a partir dos retornos previstos pelo modelo de *machine learning*, não tendo como insumo nenhum dado histórico realizado.

Foram efetuados três tipos de previsões: 3, 4 e 5 dias. Cada janela de previsão consiste em valores e cálculos diferentes no modelo e, conseqüentemente, resultados também diferentes para os retornos previstas. A janela, para fins de programação, é a variável no *script* que orienta o modelo a efetuar previsões para datas futuras.

Cada janela considerada produz valores de retornos distintos entre si. Com isso os cálculos apresentados a seguir, foram efetuados, separadamente, para cada janela de previsão.

Primeiro, calcula-se o retorno médio esperado para um determinado ativo i , a partir das previsões de retorno calculadas pelo modelo LSTM, para o período de janeiro de 2019 a setembro de 2019:

- Retorno Médio Esperado (μ_i):

$$\mu_i = \frac{1}{n} \sum_{t=1}^n \bar{r}_{it}$$

onde: μ_i é o retorno médio esperado do ativo i para o ano de 2019, \bar{r}_{it} é o retorno previsto pelo modelo de *machine learning* do ativo i para uma data t e n é o total de retornos previstos pelo modelo de *machine learning* para o período de janeiro de 2019 a setembro de 2019, neste trabalho foi utilizado um valor de 180 para n .

Adicionalmente, calcula-se a variância e, consequentemente, o desvio padrão, a partir do retorno esperado de cada ativo:

- Variância (σ_i^2):

$$\sigma_i^2 = \frac{1}{n} \sum_{t=1}^n (\bar{r}_{it} - \mu_i)^2$$

onde: σ_i^2 é a variância de um ativo i , \bar{r}_{it} é o retorno previsto pelo modelo de *machine learning* do ativo i em uma data t , μ_i é o retorno médio esperado para o ativo i , e n é o total de retornos previstos pelo modelo de *machine learning* para o período de janeiro de 2019 a setembro de 2019. O desvio padrão (σ_i): é simplesmente a raiz quadrada da variância.

Paralelamente, efetua-se o cálculo da covariância entre dois ativos da carteira e, consequentemente, a montagem da matriz de covariância:

- Covariância entre dois ativos (σ_{ij}):

$$\sigma_{ij} = \frac{1}{n} \sum_{t=1, j \neq i}^n (\bar{r}_{it} - \mu_i)(\bar{r}_{jt} - \mu_j)$$

- Matriz de Covariância (Σ):

$$\Sigma = \begin{bmatrix} \sigma_i^2 & \dots & \sigma_{ji} \\ \vdots & \ddots & \vdots \\ \sigma_{ij} & \dots & \sigma_j^2 \end{bmatrix} = \{\sigma_{ij}\}_{i,j=1}^n$$

onde : \bar{r}_{it} e \bar{r}_{jt} são, respectivamente, os retornos previstos pelo modelo de *machine learning* para o ano de 2019 dos ativos i e j em uma data t , μ_i e μ_j são, respectivamente, são os retornos médios esperados, para o período de janeiro de 2019 a setembro de 2019, para os ativos i e j e n é o total de retornos previstos pelo modelo de *machine learning* para o ano de 2019.

Finalmente, é possível formalizar o cálculo do retorno e risco (representado pelo desvio padrão) do portfólio:

- Retorno do portfólio (μ_p):

$$\mu_p = \sum_{i=1}^n \mu_i x_i = \omega^T \mu$$

- Risco do portfólio (σ_p^2):

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} = \omega^T \Sigma \omega$$

onde: σ_{ij} é a covariância entre os ativos i e j , μ_i é o retorno médio esperado para o ativo i , x_i e x_j são os percentuais alocados nos ativos i e j (a serem determinados pela otimização, descrita mais adiante), assim como ω que representa o resultado do vetor decisão de alocação (também a ser determinado pela otimização) e n é o total de retornos previstos pelo modelo de *machine learning* para o período de janeiro de 2019 a setembro de 2019, neste trabalho foi utilizado um valor de 180 para n .

Em sua primeira abordagem, Markowitz propôs um modelo que consistia em encontrar o melhor portfólio, ou seja, aquele que obtinha o melhor retorno, dada uma certa variância. Alternativamente, buscava o portfólio que apresentava a menor variância dado um certo valor aceitável para o retorno.

Naturalmente o que se deseja, dentro deste contexto, ao invés de apenas maximizar os retornos previsto, é maximizar o retorno por unidade de risco, ou seja, maximizar o Índice de Sharpe, calculados a partir dos retornos e desvios-padrão previstos pelo modelo de redes neurais.

Formalmente, as duas formulações utilizadas no trabalho são descritas a seguir:

- Minimização de Variância (v_c):

$$\begin{aligned}
 & \text{Minimizar} \sum_{i=1}^N x_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{j=1}^N x_i x_j \sigma_{ij} \\
 & \text{Sujeitos a} \sum_{i=1}^N \mu_i x_i > r_c \\
 & \sum_{i=1}^N x_i = 1 \\
 & x_i \geq 0 \quad \forall i
 \end{aligned}$$

onde: μ_i é o retorno médio esperado do ativo i , x_i é a porcentagem de alocação da carteira dada ao ativo i , x_j é a porcentagem de alocação da carteira dada ao ativo j , σ_i^2 é a variância de um ativo i , σ_{ij} é a covariância entre os ativos i e j e r_c é o retorno mínimo desejado. Não foram incluídas, neste trabalho, restrições de retorno mínimo desejado.

- Maximização de Índice de Sharpe (IS_p):

Queremos escolher o vetor de pesos da carteira, ω , de forma a maximizar a função:

$$\begin{aligned}
 \text{Maximizar } IS_p &= \frac{\omega^T \mu - R_f}{\sqrt{\omega^T \Sigma \omega}} \\
 \underline{\omega}_i &\geq 0 \quad \forall i, \\
 \overline{\omega}_i &\leq 1 \quad \forall i,
 \end{aligned}$$

onde: $\underline{\omega}_i$ são os limites mínimos e $\overline{\omega}_i$ são os limites máximos para cada ativo i , $\omega^T \mu$ é o retorno esperado do portfólio, Σ é a matriz de covariância e R_f é a taxa de retorno do ativo livre de risco.

Para haver um maior controle de exposição, ou seja, para evitar que um ativo componha grande parte da carteira, pode-se determinar um limite de exposição para os ativos. Por exemplo, limitar um ativo a menos de 25% do valor financeiro da carteira. Neste caso, estes tipos de limitações impostas pelo analista foram descartadas por não serem o objetivo da

pesquisa. Foi utilizado como retorno do ativo livre de risco a taxa de 5% ao ano (meta da taxa SELIC em dezembro de 2019).

Tendo em mente que nosso processo de tomada de decisão é eficiente, vamos apresentar as tabelas com as alocações oriundas do modelo de Markowitz. Foram construídos um total de seis portfólios de dois tipos: três destes, possuem o objetivo de se obter a menor volatilidade possível e três destes, possuem o objetivo de se obter o maior Índice de Sharpe possível.

CAPÍTULO 3 METODOLOGIA E AMOSTRA

Neste capítulo é detalhada a metodologia utilizada no trabalho, sendo descritas as etapas necessárias para concretização do objetivo desta dissertação, a construção de portfólios eficientes utilizando as técnicas propostas.

Neste trabalho procuramos examinar a aplicação do modelo proposto e avaliar o seu desempenho fora da amostra, que são os dados não vistos pelo modelo de aprendizagem, que neste contexto corresponde aos retornos do período de janeiro de 2019 a setembro de 2019, utilizando uma carteira composta por ações que fazem parte do Ibovespa. Estes portfólios foram criados a partir dos retornos das ações, de índices de análise técnica, do modelo de *machine learning* LSTM e da técnica de média variância proposta por Markowitz.

Após montadas as 6 carteiras, não foi feito nenhum tipo de rebalanceamento de carteira. Este método conhecido como *buy and hold*, foi adotado pela menor incidência de custos de transação, taxas e impostos. Nesta técnica a carteira é fixa desde sua criação e assim sofre menor interferência de fatores externos ao modelo proposto. Esta metodologia foi motivada e testada nos trabalhos de Barber e Odean (2000), Dai et al. (2008) e Donohue e Yip (2003). Todos os autores compararam carteiras com técnicas ativas de rebalanceamento diversas e o *buy and hold* e foi constatado a maior eficiência dos portfólios que utilizam esta última metodologia

3.1 SELEÇÃO DA AMOSTRA E BASE DE DADOS

A amostra consiste nos dados das 50 primeiras empresas do Ibovespa, ordenadas por capitalização de mercado, para a composição do portfólio. Estas empresas foram escolhidas de maneira arbitrária conforme a informação disponibilizada no último dia útil do mês de setembro de 2019. Através da plataforma Quantum Axis, foi possível a extração dos dados de preços de abertura, fechamento, médio, todos estes três ajustados para proventos, e volume de negociação, para as datas iniciais até o último dia útil de setembro de 2019, para todos os ativos.

Para cada ação, foi construído, através de scripts em Python, um total de 53 variáveis de análise técnica, oriundas dos indicadores e índices elencados na sessão 2.1.1, que através da análise do modelo teriam o intuito de apoiar o processo de previsão para o preço da ação no futuro.

As ações escolhidas representam cerca de 90,38% do total do Ibovespa, por capitalização de mercado, em 30/09/2019. No Capítulo 6 (Anexos) consta a relação das ações incluídas neste trabalho.

A base de dados foi dividida em duas partes, dados de treinamento e dados de teste, este método é denominado teste fora da amostra. James et al. (2013) define que dados de treino são aqueles que serão apresentados ao algoritmo de *machine learning* para criação do modelo. Os dados de treinamento são utilizados para que o algoritmo consiga compreender os padrões na base dados (os dados de teste não são acessíveis ao modelo nesta etapa) e em um problema de aprendizado supervisionado como este, cada observação consiste de uma variável de saída (retornos de preços de ações) e variáveis de entrada (indicadores de análise técnica). Dentro deste contexto, os dados de treino são os valores de retornos, volume e indicadores técnicos passados, do período entre janeiro de 2008 a dezembro de 2018.

Já os dados de teste são a parte dos dados utilizada para avaliar a performance do modelo. É importante destacar que nenhum dado da base de dados de treinamento é acessível nesta etapa. Evidentemente, neste contexto, os dados utilizados para teste são os dados de janeiro de 2019 a setembro de 2019, com exceção dos dados de retorno para as ações, pois estes serão os valores a serem previstos. Também se ressalta que neste ponto, o único dado separado para teste são os valores de preço de fechamento de cada ação.

Dada as diferentes grandezas dos dados trabalhados (preços, volume e os índices de análise técnica), antes de fornecidos ao modelo de *machine learning*, os dados são normalizados, apesar de não haver uma regra clara sobre a obrigatoriedade da utilização de normalização. Hou et al (2019) versam sobre desempenho de redes LSTM e concluem que técnicas de normalização aumentam tanto o poder de aprendizagem da rede e performance da rede. A mesma conclusão foi obtida por Laurent et al. (2016) onde comparou curvas de aprendizados da rede LSTM com e sem normalização e os resultados para os dados com diferentes tipos de normalização foram sempre mais satisfatórios que os sem normalização. Após esta etapa, efetua-se previsões para os retornos dos ativos escolhidos na base de teste para o período de janeiro de 2019 a setembro de 2019, um total de 180 retornos. Enfatiza-se o fato que o modelo de *machine learning* não tem, em sua fase de aprendizagem, contato com nenhum dado referente ao período de 2019.

A tabela 1 elenca a amostra de ações escolhida para esta pesquisa. Todas as ações fazem parte do Ibovespa, na data de 30 de setembro de 2019, e somam cerca de 90,35% do total do índice.

Tabela 1 - Ações escolhidas para seleção de portfólio.

Empresa	Código	Partição no Ibovespa em 30/09/2019 (%)
Itaú Unibanco Holding SA	ITUB4	9,173537
Vale SA	VALE3	8,30659
Banco Bradesco SA	BBDC4	7,2614
Petróleo Brasileiro SA	PETR4	6,88221
Petróleo Brasileiro SA	PETR3	4,915168
BM&FBovespa SA	B3SA3	4,86974
Ambev SA	ABEV3	4,621439
Banco do Brasil SA	BBAS3	3,350864
Itaúsa - Investimentos Itaú SA	ITSA4	3,273417
JBS SA	JBSS3	2,936836
Lojas Renner SA	LREN3	2,179456
Banco Bradesco SA	BBDC3	1,938227
IRB Brasil Resseguros S/A	IRBR3	1,938183
BRF - Brasil Foods SA	BRFS3	1,712999
All América Latina Logística SA	RAIL3	1,426773
Localiza Rent a Car SA	RENT3	1,409033
Suzano Papel e Celulose SA	SUZB3	1,349856
BB Seguridade Participacoes SA	BBSE3	1,300174
Telefonica Brasil SA	VIVT4	1,260447
Magazine Luiza SA	MGLU3	1,134424
Ultrapar Participações SA	UGPA3	1,10887
Petrobras Distribuidora SA	BRDT3	1,107802
CCR SA	CCRO3	1,063616
Weg SA	WEGE3	0,991844
Sabesp	SBSP3	0,931941
Cogna Educacao	COGN3	0,925808
Banco Santander Brasil SA	SANB11	0,907439
Azul SA	AZUL4	0,870952
Cia. Energética de Minas Gerais	CMIG4	0,767969
Lojas Americanas SA	LAME4	0,75899
Gerdau SA	GGBR4	0,746355
Pão de Açúcar	PCAR4	0,689732
BR Malls Participações SA	BRML3	0,668995
Natura Cosméticos SA	NATU3	0,655117
Centrais Elétricas Brasileiras SA	ELET3	0,598345
Estacio Participações SA	YDUQ3	0,597526
Klabin SA	KLBN11	0,542444
Centrais Elétricas Brasileiras SA	ELET6	0,524743
B2W Companhia Global do Varejo	BTOW3	0,514498
Globex Utilidades SA	VVAR3	0,502217

Cielo SA	CIEL3	0,494461
Braskem SA	BRKM5	0,475602
Cia. Siderúrgica Nacional SA	CSNA3	0,469685
Cosan SA Indústria e Comércio	CSAN3	0,456905
Multiplan	MULT3	0,430933
Bradespar SA	BRAP4	0,387394
MRV Engenharia e Participações SA	MRVE3	0,285551
Gol Linhas Aéreas Inteligentes SA	GOLL4	0,240294
Usinas Siderúrgicas de Minas Gerais S	USIM5	0,221693
Metalúrgica Gerdau SA	GOAU4	0,206242

3.2 EXPERIMENTO

Nesta sessão será apresentado um breve resumo, com a finalidade da replicação por parte do leitor do experimento efetuado neste trabalho.

- Extração da base de dados: obtenção dos preços de fechamento, abertura, máximo e mínimo (diários e ajustados para dividendos) e volume diário, de 50 ações do índice Ibovespa (vide sessão 3.1), no período de janeiro de 2008 a setembro de 2019.
- Análise técnica: elaboração dos indicadores apresentados na seção 2.2.2, para todas as 50 ações e o período janeiro de 2008 a dezembro de 2018.
- Elaboração do modelo: Foram construídos três modelos distintos para a previsão de retorno das ações. Estes modelos diferem entre si pelas janelas de previsão utilizadas: 3, 4 e 5 dias. A base de dados foi dividida em duas partes: base de treinamento e base de teste. A base de treinamento contém dados históricos de preços, volumes e indicadores técnicos de janeiro de 2008 a dezembro de 2018 e a base de teste contém os dados de volumes e indicadores de técnicos de janeiro de 2019 a setembro de 2019. A base de treinamento foi utilizada para a construção do modelo. A base de testes possui os dados de preço de janeiro de 2019 a setembro de 2019. A base de dados é normalizada, conforme explicado no item 2.2.3.

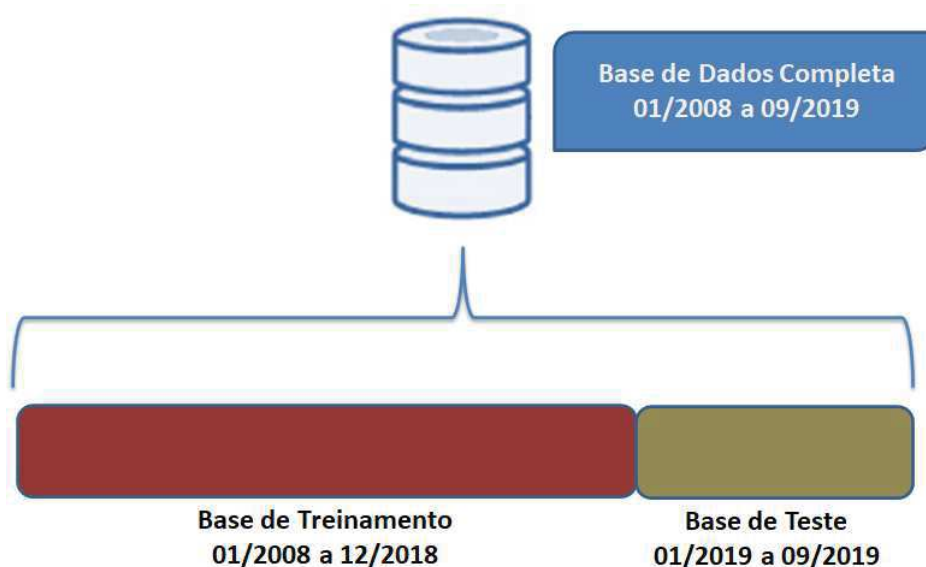


Figura 3 – Ilustração da divisão da base de dados

- Previsões: Nesta etapa, após a elaboração dos três modelos, as previsões são feitas com a base que consiste nos dados de análise técnica construídos iterativamente utilizando as previsões feitas pelo próprio modelo. O modelo fornece as previsões de preços para o período de janeiro de 2019 a setembro de 2019 (percebe-se que nesta base os valores a serem previstos, os preços não são fornecidos aos modelos de previsão). Exemplificando: Utilizando-se como exemplo o modelo que efetua a previsão para cinco períodos a frente, o dia 28/12/2018 como referencial e de posse de dados de análise técnica históricos obtidos com os preços já conhecidos, o modelo efetua a previsão para os dias 2, 3, 4, 7 e 8 de janeiro de 2019 (dado que não houve pregão no dia 29 de dezembro de 2018 ao dia 01 de janeiro de 2019 e nos dias 5 e 6 de janeiro de 2019, deste período específico). Após esta primeira etapa de previsão, o modelo calcula os dados de análise técnica utilizando no cálculo os cinco valores previstos de forma a complementar a série histórica já conhecida (ou seja, construir dados de análise técnica para os dias 2, 3, 4, 7 e 8 de janeiro de 2019) e assim, consequentemente, efetua previsões para os próximos cinco períodos. Este processo iterativo é repetido até que se chegue ao final do mês de setembro de 2019, conforme anteriormente elucidado. Analogamente, a mesma mecânica é aplicada às previsões de três e quatro períodos.

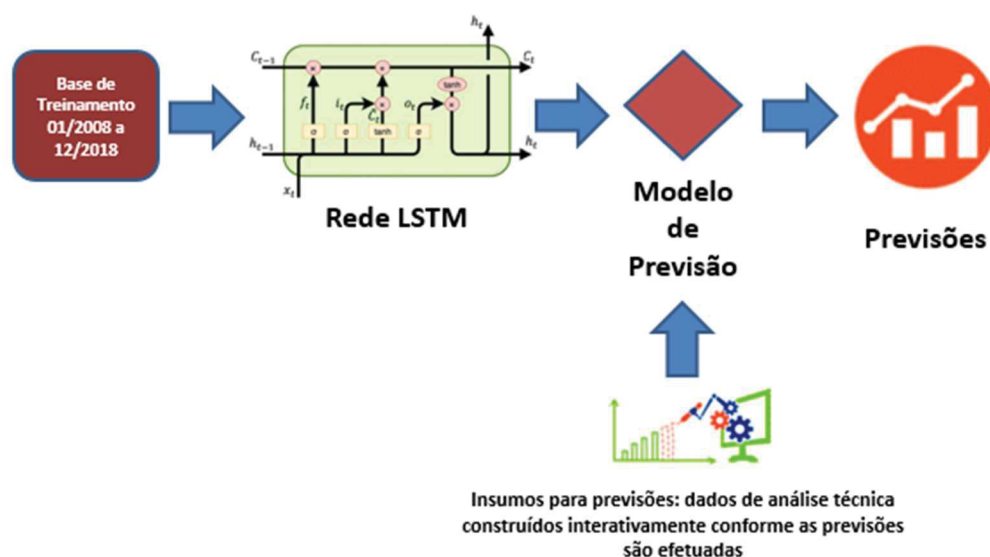


Figura 4 – Ilustração do modelo de previsão e seu funcionamento.

- **Otimização:** Com os retornos previstos pelo modelo, para o período de janeiro a setembro de 2019, conforme sessão 2.3.2, calcula-se as métricas necessárias para a otimização dos portfólios. A partir dos retornos previstos, para o período de janeiro de 2019 a setembro de 2019, extrai-se o desvio padrão desta série de retornos prevista e, finalmente, através do processo de otimização, os percentuais de alocação em cada ativo em cada carteira são calculados. O problema consiste em maximizar o índice de Sharpe e minimizar a volatilidade, esta última, dada pelo desvio padrão. Como temos três janelas de previsões distintas, são gerados seis portfólios distintos, três maximizando o índice de Sharpe e três minimizando a volatilidade.
- **Resultados:** A análise do sucesso do trabalho se deu pela comparação do índice de Sharpe e da volatilidade. É feito a verificação de desempenho analisando a performance, sobre estes atributos, no ano de 2019, dos portfólios contra o Ibovespa.

CAPÍTULO 4 RESULTADOS

Este capítulo irá apresentar os resultados dos portfólios construídos com os modelos de redes neurais e analisar a relação risco retorno destes com o portfólio de mercado (Ibovespa). Para fins de categorização, refere-se aos portfólios que fizeram a minimização da volatilidade como “Mínimo Volatilidade” ou “Min Vol” e os que maximizaram o índice Sharpe como “Máximo Sharpe” ou “Max Sharpe”. Além disso, para facilitar a visualização nos gráficos de desempenho, a nomenclatura dos portfólios terá a janela de previsão utilizada para cada experimento. Por exemplo: O portfólio que minimizar a volatilidade utilizando uma janela de previsão de 3 dias se chamará “Min Vol 3d”.

Os resultados foram organizados para responder duas perguntas.

A primeira é se seria possível a integração de ferramentas como análise técnica e *machine learning* de maneira eficiente. A conclusão da proposta demonstra a possibilidade da implementação e a qualidade dos dados obtidos a partir deste tipo de mecânica, no mercado de renda variável.

A segunda questão é o valor entregue pelo estudo. Verificando as rentabilidades das carteiras construídas com as carteiras de mercado. Queremos entender se podemos utilizar este modelo para entrega de valor para um cliente final. Comparam-se as métricas de volatilidade e Índice de Sharpe para esta conclusão.

Para fins de comparabilidade foram desconsideradas quaisquer custas eventuais que poderiam incidir sobre o trabalho, como: taxas de administração, taxa de custódia, corretagem e/ou emolumentos.

4.1 ANÁLISES DOS RESULTADOS – RETORNOS

Os gráficos 2, 3, 4 e 5 referem-se aos retornos diários das carteiras elaboradas com previsões de 3, 4 e 5 dias.

Percebe-se que ao analisar apenas a variável “retorno”, as carteiras de minimização de volatilidade, possuem um desempenho semelhante à da carteira de mercado (Ibovespa). Porém, constata-se que as carteiras de maximização de Índice de Sharpe, possuem um retorno absoluto bem superior à carteira de mercado e as carteiras de minimização de volatilidade.

Ratificando com enunciado no início deste capítulo, foram desprezadas quaisquer custas para a quantificação dos retornos dos portfólios construídos e o período de comparação ao *benchmark* foi todo o ano de 2019.

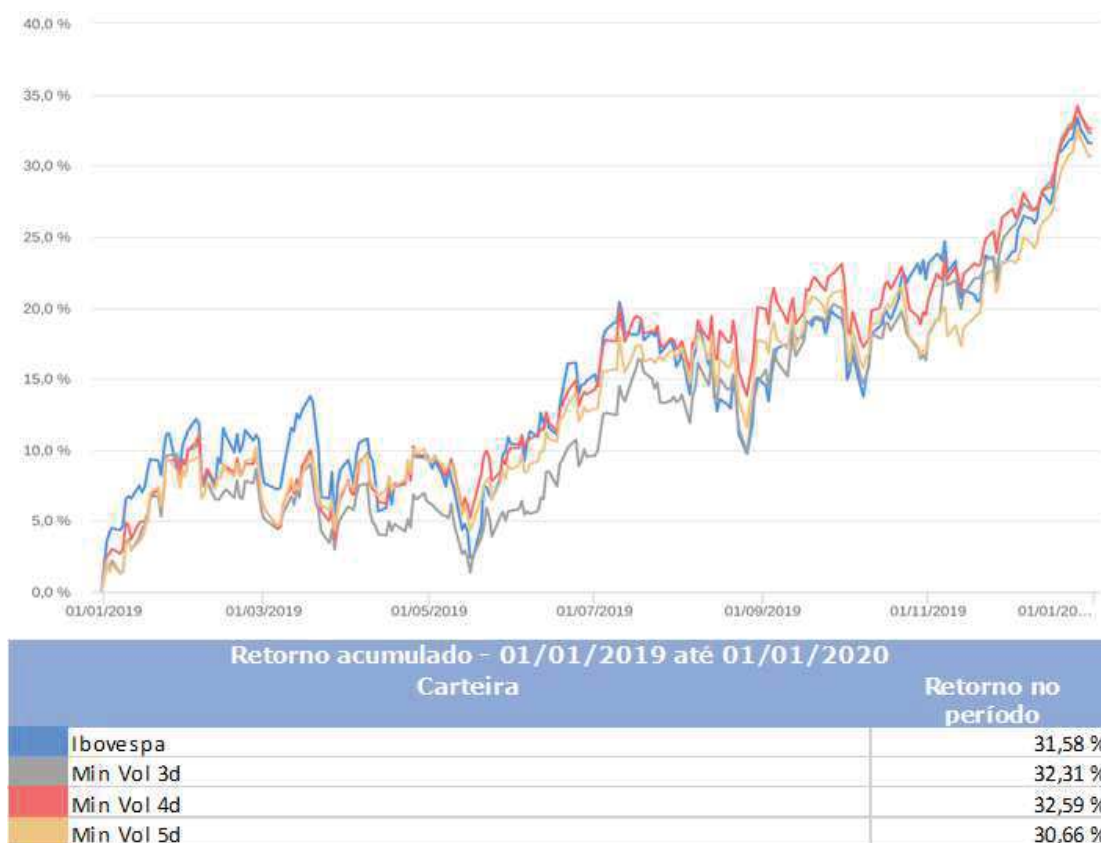


Gráfico 2 - Comparativo entre os retornos acumulados dos portfólios que efetuam a minimização da volatilidade e do Ibovespa entre os períodos de janeiro de 2019 a dezembro de 2019.



Gráfico 3: Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe e do Ibovespa entre os períodos de janeiro de 2019 a dezembro de 2019.



Retorno acumulado - 01/01/2019 até 31/08/2020	
Ativo	Retorno Acumulado no Período
Max Sharpe 3d	120,92 %
Max Sharpe 4d	68,27 %
Max Sharpe 5d	35,46 %
Ibovespa	13,06 %

Gráfico 4 - Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe e do Ibovespa entre os períodos de janeiro de 2019 a agosto de 2020.

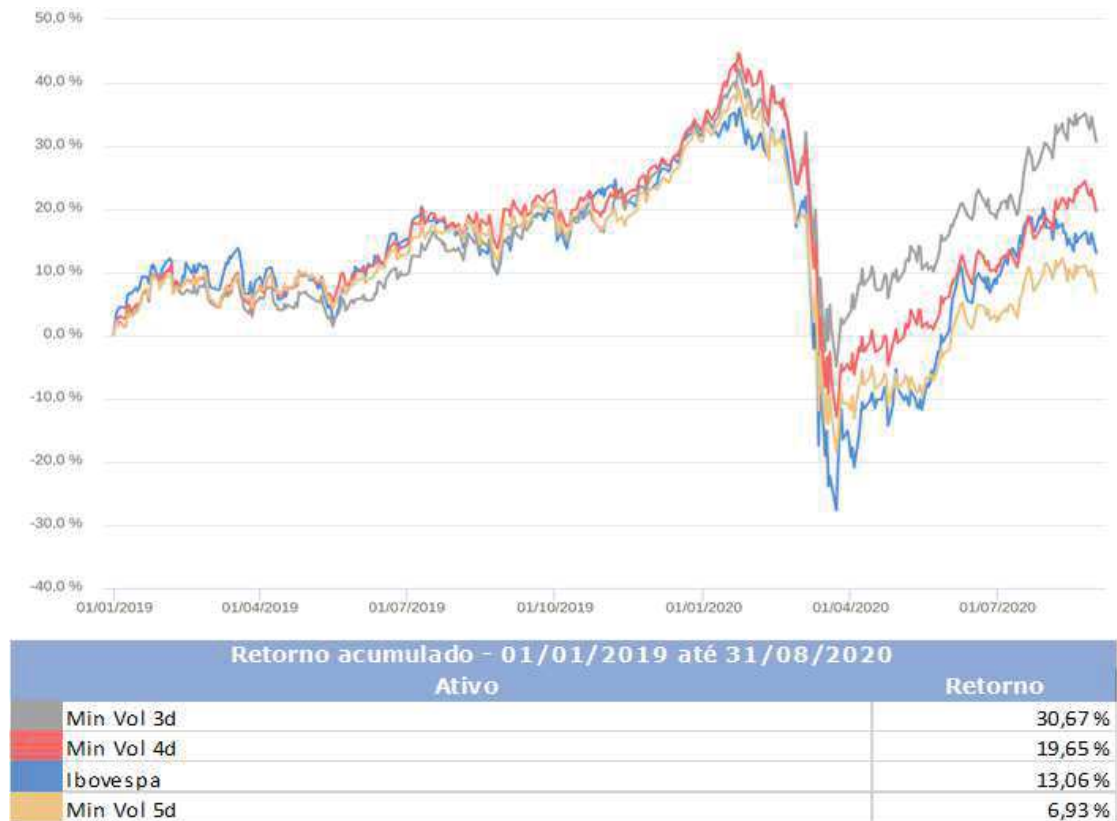
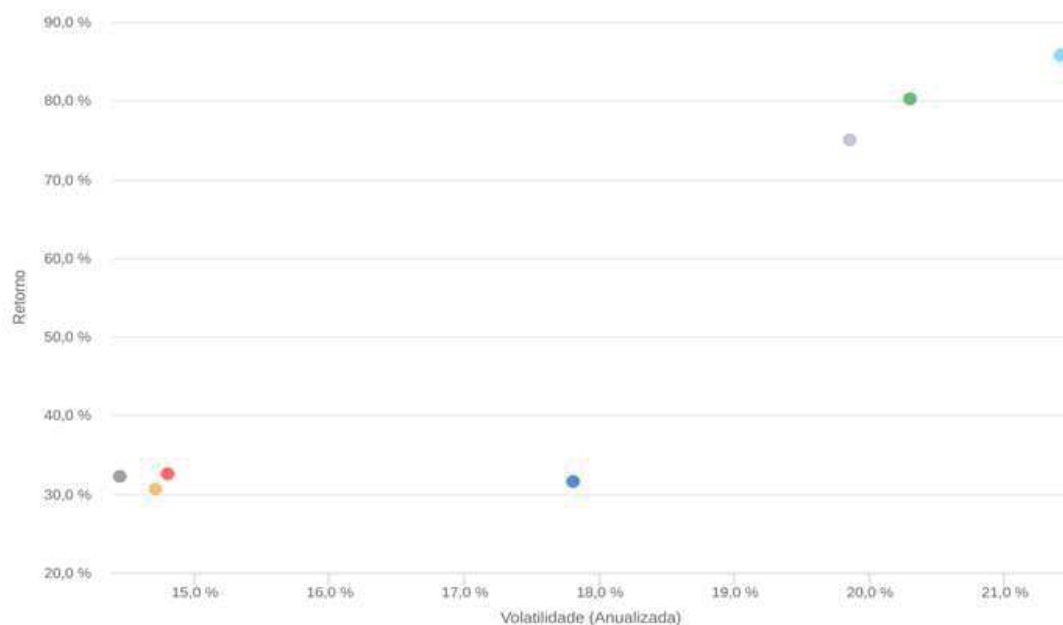


Gráfico 5 - Comparativo entre os retornos acumulados dos portfólios que efetuam a minimização da volatilidade e do Ibovespa entre os períodos de janeiro de 2019 a agosto de 2020.

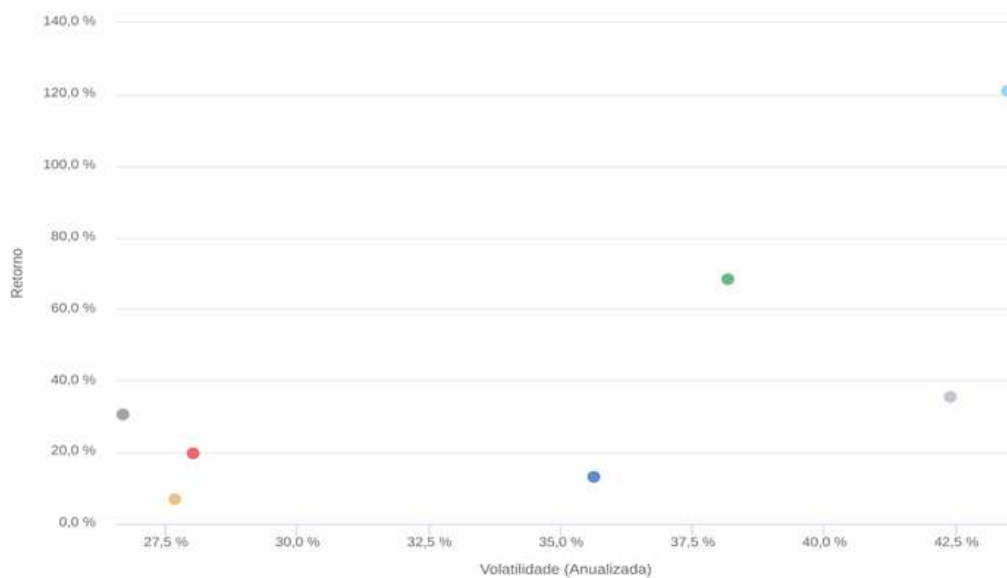
4.2 ANÁLISES DOS RESULTADOS – RISCO-RETORNO

Os gráficos 6 e 7 abaixo ilustram a relação risco e retorno. O eixo da horizontal que quantifica o risco de cada carteira é uma medida de desvio padrão anualizado.



Risco X Retorno de 01/01/2019 até 31/12/2019				
Ativo	Retorno	Volatilidade (Anualizada)	Índice de Sharpe	
Max Sharpe 3d	85,81 %	21,41 %	3,77	
Max Sharpe 4d	80,27 %	20,30 %	3,71	
Max Sharpe 5d	74,98 %	19,86 %	3,52	
Min Vol 3d	32,31 %	14,45 %	1,89	
Min Vol 4d	32,59 %	14,80 %	1,86	
Min Vol 5d	30,66 %	14,71 %	1,74	
Ibovespa	31,58 %	17,81 %	1,49	

Gráfico 6 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre o período de janeiro de 2019 a dezembro de 2019.



Risco X Retorno de 01/01/2019 até 31/08/2020				
Ativo	Retorno	Volatilidade (Anualizada)	Índice de Sharpe	
Max Sharpe 3d	120,92 %	43,48 %	2,67	
Max Sharpe 4d	68,27 %	38,18 %	1,66	
Min Vol 3d	30,67 %	26,70 %	0,96	
Max Sharpe 5d	35,46 %	42,39 %	0,72	
Min Vol 4d	19,65 %	28,02 %	0,52	
Ibovespa	13,06 %	35,64 %	0,23	
Min Vol 5d	6,93 %	27,68 %	0,07	

Gráfico 7 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre o período de janeiro de 2019 a agosto de 2020.

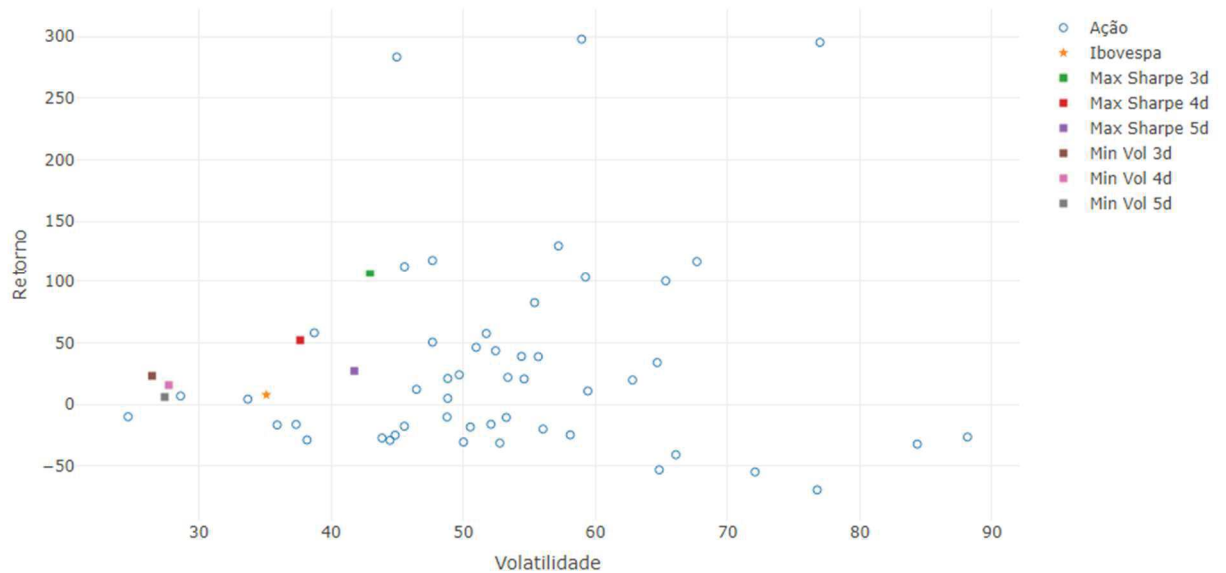


Gráfico 8 - Relação retorno acumulado e volatilidade anualizada dos portfólios e de todas as ações individuais que os compõem no período entre janeiro de 2019 a agosto de 2020

Verifica-se através da análise do Gráfico 8 que é possível inferir uma linha imaginária representando a fronteira eficiente com os portfólios em cima da linha. Também se verifica três ações na parte superior do gráfico. Dado o período da análise de janeiro de 2019 a agosto de 2020 é aceitável que alguns ativos tenham uma relação retorno-risco superior ao mercado.

Análises dos Resultados – TABELAS DE RETORNO MENSAL

A tabela abaixo traz o resumo da rentabilidade das carteiras construídas em todos os meses do ano de 2019. Também há a comparação destas com o seu respectivo *benchmark*, o Ibovespa.

Tabela 2 - Retornos mensais dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e o Ibovespa, de todos os meses do ano de 2019.

Data	Ibovespa	Min Vol 3d	Min Vol 4d	Min Vol 5d	Max Sharpe 3d	Max Sharpe 4d	Max Sharpe 5d
janeiro, 2019	10,8 %	8,9 %	9,0 %	8,1 %	14,0 %	19,2 %	13,4 %
fevereiro, 2019	-1,9 %	-3,0 %	-2,2 %	-1,4 %	-5,6 %	-5,9 %	-1,9 %
março, 2019	-0,2 %	-0,6 %	0,0 %	0,3 %	3,3 %	3,8 %	1,9 %
abril, 2019	1,0 %	1,2 %	2,6 %	2,3 %	11,6 %	10,0 %	9,7 %
maio, 2019	0,7 %	-0,6 %	0,8 %	-0,6 %	2,3 %	6,7 %	3,4 %
junho, 2019	4,1 %	3,6 %	3,4 %	3,7 %	5,0 %	2,5 %	2,4 %
julho, 2019	0,8 %	3,5 %	2,8 %	3,4 %	15,4 %	11,2 %	9,8 %
agosto, 2019	-0,7 %	1,2 %	2,5 %	1,0 %	7,3 %	8,8 %	4,9 %
setembro, 2019	3,6 %	4,6 %	2,5 %	3,0 %	2,9 %	4,2 %	4,5 %
outubro, 2019	2,4 %	-3,0 %	-2,9 %	-3,6 %	2,6 %	-4,5 %	1,0 %
novembro, 2019	0,9 %	7,5 %	5,9 %	5,4 %	3,2 %	2,9 %	3,6 %
dezembro, 2019	6,8 %	5,8 %	4,8 %	6,1 %	3,4 %	3,7 %	5,5 %

4.3 ANÁLISES DOS RESULTADOS – DECOMPOSIÇÃO DE CARTEIRAS

Nesta parte, as carteiras foram decompostas para a análise de concentração em cada setor econômico, possibilitando um maior entendimento dos resultados obtidos por cada resultado. Além disso também foi feito uma análise comparativa de concentração por setor econômico de cada portfólio contra o Ibovespa. Os dados levam em conta a data de início da análise dos portfólios bem como a composição do Ibovespa em janeiro de 2019. O padrão para classificação utilizado foi o Global Industry Classification Standard, conhecido pela sigla GICS. Esse padrão é um sistema de classificação para ações desenvolvidas conjuntamente pela Morgan Stanley Capital International (MSCI) e Standard & Poor's.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Bens de consumo discricionário			30,79	6,93	23,86		
	MGLU3	MAGAZINE LUIZA SA	24,83	0,84	23,99	317,90%	98,28%
	MRVE3	MRV ENGENHARIA	1,20	0,25	0,95	47,57%	-15,36%
	VVAR3	VIA VAREJO SA	4,76	0,12	4,64	360,59%	81,02%
Produtos básicos de consumo			24,24	9,54	14,70		
	BRFS3	BRF SA	16,62	1,21	15,41	-9,17%	-43,41%
	JBSS3	JBS SA	7,62	1,29	6,33	97,41%	-11,32%
Serviços financeiros			13,24	36,51	-23,27		
	B3SA3	B3 SA-BRASIL BOLSA BALCAO	6,65	3,79	2,86	127,60%	42,01%
	IRBR3	IRB BRASIL RESSEGUROS SA	6,59		6,59	-70,83%	-79,16%
Industriais			20,00	5,49	14,51		
	AZUL4	AZUL SA	2,00		2,00	-36,44%	-60,74%
	CCRO3	CCR SA	2,38	0,89	1,49	20,00%	-29,19%
	RAIL3	RUMO SA	14,69	1,26	13,43	39,76%	-8,97%
	WEGE3	WEG SA	0,93	0,89	0,04	283,07%	93,85%
Materiais			1,19	16,67	-15,48		
	CSNA3	CIA SIDERURGICA NACIONAL SA	1,19	0,39	0,80	72,51%	8,08%
Serviços de utilidade pública			10,54	5,05	5,49		
	SBSP3	CIA SANEAMENTO BASICO DE SP	10,54	0,70	9,84	59,30%	-17,14%

Tabela 3 – Decomposição do portfólio Max Sharpe 3d por percentual alocado em cada setor e ativo do portfólio e do benchmark e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Serviços de comunicação			11,48	2,13	9,35		
	VVT4	TELEFONICA BRASIL S.A.-PREF	11,48	1,31	10,17	7,18%	-14,54%
Bens de consumo discricionário			16,13	6,93	9,20		
	MGLU3	MAGAZINE LUIZA SA	4,34	0,84	3,50	317,90%	98,28%
	MRVE3	MRV ENGENHARIA	9,04	0,25	8,79	47,57%	-15,36%
	VVAR3	VIA VAREJO SA	2,75	0,12	2,63	360,59%	81,02%
Produtos básicos de consumo			29,40	9,54	19,86		
	BRFS3	BRF SA	9,71	1,21	8,50	-9,17%	-43,41%
	JBSS3	JBS SA	19,69	1,29	18,40	97,41%	-11,32%
Energia			9,67	14,02	-4,35		
	CSAN3	COSAN SA	9,67	0,37	9,30	156,87%	23,54%
Serviços financeiros			10,32	36,51	-26,19		
	B3SA3	B3 SA-BRASIL BOLSA BALCAO	0,14	3,79	-3,65	127,60%	42,01%
	IRBR3	IRB BRASIL RESSEGUROS SA	10,18		10,18	-70,83%	-79,16%
Industriais			18,05	5,49	12,56		
	AZUL4	AZUL SA	1,13		1,13	-36,44%	-60,74%
	CCRO3	CCR SA	5,55	0,89	4,66	20,00%	-29,19%
	RAIL3	RUMO SA	9,75	1,26	8,49	39,76%	-8,97%
	WEGE3	WEG SA	1,62	0,89	0,73	283,07%	93,85%
Materiais			0,44	16,67	-16,23		
	CSNA3	CIA SIDERURGICA NACIONAL SA	0,44	0,39	0,05	72,51%	8,08%
Serviços de utilidade pública			4,51	5,05	-0,54		
	ELET6	CENTRAIS ELETRICAS BRAS-PR B	4,51	0,49	4,02	33,33%	-1,78%

Tabela 4 - Decomposição do portfólio Max Sharpe 4d por percentual alocado em cada setor e ativo do portfólio e do *benchmark* e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Serviços de comunicação			1,96	2,13	-0,17		
	VIVT4	TELEFONICA BRASIL S.A.-PREF	1,96	1,31	0,65	7,18%	-14,54%
Bens de consumo discricionário			22,76	6,93	15,83		
	COGN3	COGNA EDUCACAO	9,93	0,91	9,02	-32,24%	-47,42%
	MGLU3	MAGAZINE LUIZA SA	12,83	0,84	11,99	317,90%	98,28%
Produtos básicos de consumo			21,41	9,54	11,87		
	BRFS3	BRF SA	17,73	1,21	16,52	-9,17%	-43,41%
	NTCO3	NATURA &CO HOLDING SA	3,68		3,68	124,11%	30,40%
Energia			19,04	14,02	5,02		
	CSAN3	COSAN SA	19,04	0,37	18,67	156,87%	23,54%
Serviços financeiros			25,13	36,51	-11,38		
	B3SA3	B3 SA-BRASIL BOLSA BALCAO	5,04	3,79	1,25	127,60%	42,01%
	IRBR3	IRB BRASIL RESSEGUROS SA	20,09		20,09	-70,83%	-79,16%
Industriais			8,49	5,49	3,00		
	CCRO3	CCR SA	1,14	0,89	0,25	20,00%	-29,19%
	RAIL3	RUMO SA	7,35	1,26	6,09	39,76%	-8,97%
Materiais			1,21	16,67	-15,46		
	CSNA3	CIA SIDERURGICA NACIONAL SA	1,21	0,39	0,82	72,51%	8,08%

Tabela 5 - Decomposição do portfólio Max Sharpe 5d por percentual alocado em cada setor e ativo do portfólio e do *benchmark* e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Serviços de comunicação			37,05	2,13	34,92		
	VIVT4	TELEFONICA BRASIL S.A.-PREF	37,05	1,31	35,74	7,18%	-14,54%
Bens de consumo discricionário			3,30	6,93	-3,63		
	BRDT3	PETROBRAS DISTRIBUIDORA SA	3,30		3,30	-14,51%	-26,94%
Produtos básicos de consumo			12,43	9,54	2,89		
	BRFS3	BRF SA	4,92	1,21	3,71	-9,17%	-43,41%
	PCAR4	CIA BRASILEIRA DE DIS-PREF	1,23	0,86	0,37		
	NTCO3	NATURA &CO HOLDING SA	6,28		6,28	124,11%	30,40%
Energia			5,22	14,02	-8,80		
	UGPA3	ULTRAPAR PARTICIPACOES SA	5,22	1,96	3,26	-23,76%	-20,41%
Serviços financeiros			8,92	36,51	-27,59		
	BBSE3	BB SEGURIDADE PARTICIPACOES	7,41	1,27	6,14	-0,14%	-26,92%
	ITUB4	ITAU UNIBANCO HOLDING S-PREF	1,51	11,69	-10,18	-31,18%	-34,15%
Industriais			8,24	5,49	2,75		
	RAIL3	RUMO SA	5,87	1,26	4,61	39,76%	-8,97%
	WEGE3	WEG SA	2,37	0,89	1,48	283,07%	93,85%
Materiais			24,84	16,67	8,17		
	KLBN11	KLABIN SA - UNIT	6,05	0,64	5,41	64,80%	42,15%
	SUZB3	SUZANO SA	13,68	1,21	12,47	33,90%	28,50%
	VALE3	VALE SA	5,11	10,47	-5,36	19,80%	14,63%

Tabela 6 - Decomposição do portfólio Min Vol 3d por percentual alocado em cada setor e ativo do portfólio e do *benchmark* e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Serviços de comunicação			33,18	2,13	31,05		
	VIVT4	TELEFONICA BRASIL S.A.-PREF	33,18	1,31	31,87	7,18%	-14,54%
Bens de consumo discricionário			0,26	6,93	-6,67		
	BRDT3	PETROBRAS DISTRIBUIDORA SA	0,26		0,26	-14,51%	-26,94%
Produtos básicos de consumo			9,50	9,54	-0,04		
	ABEV3	AMBEV SA	3,50	4,69	-1,19	-17,10%	-31,71%
	BRFS3	BRF SA	3,02	1,21	1,81	-9,17%	-43,41%
	JBSS3	JBS SA	2,98	1,29	1,69	97,41%	-11,32%
Energia			2,75	14,02	-11,27		
	UGPA3	ULTRAPAR PARTICIPACOES SA	2,75	1,96	0,79	-23,76%	-20,41%
Serviços financeiros			11,93	36,51	-24,58		
	SANB11	BANCO SANTANDER BRASIL-UNIT	1,11	1,14	-0,03	-31,43%	-40,87%
	IRBR3	IRB BRASIL RESSEGUROS SA	6,60		6,60	-70,83%	-79,16%
	ITUB4	ITAU UNIBANCO HOLDING S-PREF	4,22	11,69	-7,47	-31,18%	-34,15%
Industriais			7,85	5,49	2,36		
	AZUL4	AZUL SA	0,59		0,59	-36,44%	-60,74%
	RENT3	LOCALIZA RENT A CAR	7,26	1,04	6,22	76,47%	5,46%
Materiais			27,20	16,67	10,53		
	GGBR4	GERDAU SA-PREF	0,79	1,00	-0,21	32,25%	-2,00%
	KLBN11	KLABIN SA - UNIT	16,87	0,64	16,23	64,80%	42,15%
	GOAU4	METALURGICA GERDAU SA-PREF	3,81	0,27	3,54	28,82%	-3,66%
	SUZB3	SUZANO SA	3,85	1,21	2,64	33,90%	28,50%
	VALE3	VALE SA	1,88	10,47	-8,59	19,80%	14,63%
Serviços de utilidade pública			7,33	5,05	2,28		
	CMIG4	CIA ENERGETICA MINAS GER-PRF	7,33	0,93	6,40	-17,59%	-17,18%

Tabela 7 - Decomposição do portfólio Min Vol 4d por percentual alocado em cada setor e ativo do portfólio e do *benchmark* e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Setor	Código	Empresa	Participação em 01/2019 (%)			Retorno no Período (%)	
			Portfólio	Ibovespa	+/-	01/2019 - 08/2020	01/2020 - 08/2020
Serviços de comunicação			30,29	2,13	28,16		
	VIVT4	TELEFONICA BRASIL S.A.-PREF	30,29	1,31	28,98	7,18%	-14,54%
Bens de consumo discricionário			1,54	6,93	-5,39		
	COGN3	COGNA EDUCACAO	1,54	0,91	0,63	-32,24%	-47,42%
Produtos básicos de consumo			16,89	9,54	7,35		
	ABEV3	AMBEV SA	10,31	4,69	5,62	-17,10%	-31,71%
	BRFS3	BRF SA	3,46	1,21	2,25	-9,17%	-43,41%
	PCAR4	CIA BRASILEIRA DE DIS-PREF	3,12	0,86	2,26		
Energia			5,10	14,02	-8,92		
	UGPA3	ULTRAPAR PARTICIPACOES SA	5,10	1,96	3,14	-23,76%	-20,41%
Serviços financeiros			14,48	36,51	-22,03		
	BBSE3	BB SEGURIDADE PARTICIPACOES	4,88	1,27	3,61	-0,14%	-26,92%
	IRBR3	IRB BRASIL RESSEGUROS SA	9,60		9,60	-70,83%	-79,16%
Industriais			10,08	5,49	4,59		
	RENT3	LOCALIZA RENT A CAR	10,08	1,04	9,04	76,47%	5,46%
Tecnologia da informação			0,32	0,72	-0,40		
	CIEL3	CIELO SA	0,32	0,72	-0,40	-47,58%	-44,32%
Materiais			19,59	16,67	2,92		
	GGBR4	GERDAU SA-PREF	2,07	1,00	1,07	32,25%	-2,00%
	SUZB3	SUZANO SA	9,52	1,21	8,31	33,90%	28,50%
	VALE3	VALE SA	8,00	10,47	-2,47	19,80%	14,63%
Serviços de utilidade pública			1,71	5,05	-3,34		
	CMIG4	CIA ENERGETICA MINAS GER-PRF	1,71	0,93	0,78	-17,59%	-17,18%

Tabela 8 - Decomposição do portfólio Min Vol 5d por percentual alocado em cada setor e ativo do portfólio e do benchmark e o retorno acumulado de cada ativo para os períodos de Janeiro/2019 a Agosto/2020 e Janeiro/2020 a Agosto/2020.

Analisando a Tabela 3, que versa sobre o portfólio Max Sharpe 5d, verifica-se altas concentrações em três ativos: Magazine Luiza (MGLU3), BRF (BRFS3) e Rumo (RAIL3). Pode-se concluir que o retorno do montante alocado da Magazine Luiza foi responsável por 78,33% dos 120,92% de retorno acumulado (janeiro de 2019 a agosto de 2020) do portfólio. Ressalta-se o fator desse portfólio manter uma menor concentração no Setor Financeiro, em relação ao *benchmark*. Setor este bastante impactado pela crise do COVID-19 em 2020.

Analisando a Tabela 4, que versa sobre o portfólio Max Sharpe 4d, pode-se verificar cerca de 20 pontos percentuais a mais em concentração no setor de Produtos Básicos de Consumo e 13 pontos percentuais a mais em Industriais, em relação ao *benchmark*. Tal concentração, alinhada as performances notáveis de ações como Magazine Luiz, Via Varejo e WEG, pode explicar a maior performance deste portfólio em relação ao Ibovespa.

Analisando a Tabela 5, que versa sobre o portfólio Max Sharpe 5d, pode-se verificar a alta concentração no setor de Bens de Consumo Discricionário, com destaque para alta concentração em Magazine Luiza, e no setor de Serviços Financeiros, com destaque para a elevada posição detida em IRB Resseguros. Apesar do portfólio em questão ter superado seu *benchmark*, o Ibovespa, a elevada posição em IRB, que somente no ano de 2020 sofreu uma queda de 79% (janeiro 2020 a agosto de 2020) no preço de suas ações, explica o porquê deste portfólio apresentar o menor desempenho entre os portfólios que efetuam a maximização do índice de Sharpe na análise de período entre janeiro de 2019 a agosto de 2020.

Analisando a Tabela 6, que versa sobre o portfólio Min Vol 3d, verifica-se uma alta concentração em Telefônica. A empresa geração de caixa constante e bastante previsível, segundo Reis, por estar inserida em um segmento anticíclico, pouco afetado pelas crises econômicas e que possui alta recorrência de negócios. Dado este contexto entende-se como razoável que esta ação tenha uma baixa volatilidade, sendo apropriada para portfólios de baixa volatilidade. O investimento em empresas como Suzao, Klabin e Natura também explicam como o retorno apurado para este portfólio foi maior que seu *benchmark*.

Analisando a Tabela 7, que versa sobre o portfólio Min Vol 4d, também se verifica uma alta concentração em Telefônica, pelos mesmos motivos já elucidados para o portfólio Min Vol 4d. A alta concentração no setor de Materiais e a menor concentração no setor de Serviços Financeiros, em relação ao *benchmark*, pode explicar o porquê deste portfólio apresentar um retorno maior comparativamente ao Ibovespa.

Analisando a Tabela 8, que versa sobre o portfólio Min Vol 5d, assim como os portfólios Min Vol 3d e Min Vol 4d, verifica-se uma alta concentração em ações da Telefônica.

Investimentos alinhados à Telefônica como IRB e AMBEV podem explicar os motivos deste portfólio apresentar um pior desempenho em relação ao Ibovespa.

CAPÍTULO 5 CONCLUSÃO

Este trabalho tem como objetivo analisar a utilização de *Machine Learning* e indicadores de análise técnica em colaboração para elaboração de portfólios, em uma carteira composta exclusivamente de ações. A análise é feita a partir de uma amostra que compreende as 50 ações do Ibovespa com maior capitalização em 30/10/2019. A metodologia é semelhante a estudos encontrados, por exemplo, Obeidat et al. (2018) e Chaves et al. (2010), onde modelos de otimização foram utilizados para comparar o retorno total de portfólio de ações. Há uma pequena quantidade de trabalhos, em âmbito nacional, com este viés.

Utilizamos o algoritmo de rede neural recorrente LSTM (*Long Short Term Memory*) para prever os retornos das ações, para três, quatro e cinco dias, e construir sua matriz de covariância. Os *inputs* do modelo são indicadores de análise técnica e os próprios preços das ações.

Após a obtenção dos retornos pelo modelo, aplica-se o modelo de média variância de Markowitz a fim de se obter, para cada janela de previsão, o portfólio de maior Índice Sharpe e o portfólio de menor volatilidade. O método *Buy And Hold* foi utilizado para fazer comparar das carteiras construídas com o *benchmark* adequado, ou seja, as carteiras não foram rebalanceadas durante o período fora da amostra.

A contribuição deste trabalho não é só a de estimular o uso de *machine learning* e redes neurais no mercado de capitais, mas de contribuir com um modelo que possa enriquecer a tomada de decisão para gestores e investidores de renda variável. Esse trabalho propõe algo além, que é a aplicação da *machine learning*, análise técnica e otimização de portfólio para carteiras exclusivamente compostas por ativos de renda variável, entregando uma ferramenta completa, pronta para implementação. Desta forma, pode-se demonstrar como diferentes ferramentas podem, em conjunto, contribuir para o fornecimento de carteiras de investimento rentáveis muito com a prática profissional.

Dado o modelo construído faz-se interessante, para estudos posteriores, a comparação desta modelagem com *benchmarks* tradicionais como *Random Walk* e o modelo ARMA.

Constata-se que as carteiras construídas obtiveram bons resultados. Os portfólios que maximizam o Índice de Sharpe obtiveram uma relação retorno x risco bem superior ao benchmark e as carteiras cujo objetivo era minimizar a volatilidade conseguiram se manter indexados ao *benchmark* com uma volatilidade menor.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi. **The Economics of Artificial Intelligence: An Agenda** (National Bureau of Economic Research Conference Report). [S. l.]: University of Chicago Press, 2019. 648 p.

ARMS, Richard W. **Trading without fear: eliminating emotional decisions with arms trading strategies**. John Wiley & Sons, 1996.

ASSAF NETO, Alexandre. **Mercado Financeiro**. 5ªed. São Paulo: Atlas, 2003.

BADOLA, Akshay; CHERIAN, Thomas; PADMANABHAN, Vineet. Multi-cell LSTM Based Neural Language Model. **arXiv preprint arXiv:1811.06477**, 2018.

BARBER, B.M. and ODEAN, T., Trading is hazardous to your wealth: the common stock investment performance of individual investors. *J. Finan.*, 2000, 55, 773–806

BARBOSA, Marcos Joaquim. **Análise gráfica produz boas rentabilidades?: uma avaliação da eficácia da análise técnica computadorizada na geração de retornos**. 2007. Dissertação de Mestrado. Universidade Federal de Pernambuco.

BLAU, William. True strength index. **Tech Anal Stocks Commod (traders. com)**, v. 11, n. 1, p. 438-446, 1991.

BOAINAIN, Pedro Gabriel. “Ombro-cabeça-ombro”: Testando a lucratividade do padrão gráfico de análise técnica no mercado de ações brasileiro. 2007.

BOLLEN, Nicolas P. B. Mutual Fund Attributes and Investor Behavior. **Journal of Financial and Quantitative Analysis**. v. 42, n. 3, p. 683-708, 2007

BOLLINGER, John. **Bollinger on Bollinger bands**. McGraw Hill Professional, 2002.

CALVO, Clara *et al.* Fuzzy portfolio selection with non-financial goals: exploring the efficient frontier. **Annals of Operations Research**, v. 245, p. 31-46, 2016.

DE CAMARGOS, Marcos Antônio; BARBOSA, Francisco Vidal. Análise empírica da reação do mercado de capitais brasileiro aos anúncios de fusões e aquisições ocorridos entre 1994 e 2001. **Revista de Administração**, v. 42, n. 4, p. 468-481, 2007.

CASHMAN, George *et al.* Investor behavior in the mutual fund industry: evidence from gross flows. **Journal of Economics and Finance**, v. 38, p. 541-567, 2014.

CERVELLÓ-ROYO, Roberto; GUIJARRO, Francisco; MICHNIUK, Karolina. : Forecasting the DJIA index with intraday data. **Expert systems with Applications**, v. 42, n. 14, p. 5963-5975, 2015.

CHACKO, Ashish; GANGULY, Shruti; RAVICHANDRA, Thangjam. TECHNICAL ANALYSIS OF EQUITY SHARES. **International Journal of Research in Finance & Marketing**, p. 10-17, 2015.

CHAVES, Denis et al. Risk parity portfolio vs. other asset allocation heuristic portfolios. **The Journal of Investing**, v. 20, n. 1, p. 108-118, 2011.

CHEN, Shu-Heng; WANG, Paul P. **Computational Intelligence in Economics and Finance**. 2004.

CHENG, Haibin et al. Multistep-ahead time series prediction. In: **Pacific-Asia Conference on Knowledge Discovery and Data Mining**. Springer, Berlin, Heidelberg, 2006. p. 765-774.

CHEW, Boon-Kiat; WONG, Wing-Keung; MANZUR, Meher. How rewarding is technical analysis? Evidence from Singapore stock market. **Applied Financial Economics**, v. 13, n. 7, p. 543-551, 2003.

CHICAROLI, Rodrigo; VALLS PEREIRA, Pedro L. Predictability of equity models. **Journal of Forecasting**, v. 34, n. 6, p. 427-440, 2015.

CHONG, Terence Tai-Leung; NG, Wing-Kam. Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. **Applied Economics Letters**, v. 15, n. 14, p. 1111-1114, 2008.

CHOUDHRY, Rohit; GARG, Kumkum. **A hybrid machine learning system for stock market forecasting**. World Academy of Science, Engineering and Technology, v. 39, n. 3, p. 315-318, 2008.

CHOURMOUZIADIS, Konstandinos; CHATZOGLOU, Prodromos D. An intelligent short term stock trading fuzzy system for assisting investors in portfolio management. **Expert Systems with Applications**, v. 43, p. 298-311, 2016.

COVEL, Michael. **Trend following (updated edition): Learn to make millions in up or down markets**. FT Press, 2009.

CUI, Lingxiao; LONG, Wen; LU, Zhichen. Deep learning-based feature engineering for stock price movement prediction. **Knowledge-Based Systems**, v. 164, p. 163-173, 2019.

CYERT, R.; MARCH, J. **A behavioral theory of the firm**. Englewood Cliffs: Prentice-Hall, 1963.

DASH, Rajashree; DASH, Pradipta Kishore. **A hybrid stock trading framework integrating technical analysis with machine learning techniques**. The Journal of Finance and Data Science, v. 2, n. 1, p. 42-57, 2016.

DE CARVALHO, Luis Gustavo Pinto; DA COSTA JR, Newton CA; DE OV GOULART, Marco Antônio. Análise técnica versus hipótese dos mercados eficientes: um estudo utilizando o indicador MACD. **Revista Alcance**, v. 15, n. 3, p. 398-416, 2008.

DE FREITAS, FABIO D.; DE ALMEIDA, AILSON R. Portfolio selection with predicted returns using neural networks. In: **IASTED International Conference on Artificial Intelligence and Applications**. 2001. p. 99-103.

DELONG, J.; SHLEIFER, A.; SUMMERS, L.; WALDMANN, R. Noise Trader Risk in Financial Markets. **Journal of Political Economy**, [S. l.], n. 98(4), p. 703-738, 1990.

DONG, Bing; LI, Zhaoxuan. Short term predictions of occupancy in commercial buildings: Performance analysis for stochastic models and machine learning approaches. **Energy and Buildings**, [s. l.], 2018

DONOHUE, Christopher; YIP, Kenneth. Optimal portfolio rebalancing with transaction costs. **The Journal of Portfolio Management**, v. 29, n. 4, p. 49-63, 2003.

DORSEY, Donald. The Mass Index by Donald Dorsey. **Technical Analysis of Stocks & Commodities**, [s. l.], 1992.

DUAN, Yanqing; EDWARDS, John S.; DWIVEDI, Yogesh K. **Artificial intelligence for decision making in the era of Big Data**: evolution, challenges and research agenda. *International Journal of Information Management*, ed. 48, p. 63-71, 2019.

ELDER, Alexander. **Trading for a living: psychology, trading tactics, money management**. John Wiley & Sons, 1993.

FAMA, E. F. Efficient Capital Markets: a review of theory and empirical work. **The Journal of Finance**, v. 25, n. 2, p. 383-417, 1970.

FANT, L. Franklin; O'NEAL, Edward. Temporal changes in the determinants of mutual fund flows. **Journal of Financial Research**, v. 23, p. 353–371, 2000.

FINKELSTEIN, Joseph; JEONG, In Cheol. Machine learning approaches to personalize early prediction of asthma exacerbations. **Annals of the New York Academy of Sciences**, [s. l.], ed. 1387, 2016.

FRÄNKEL, Constantin. **Long short-term memory neural networks for one-step time series forecasting**. 2018. Tese de Doutorado. Wien.

JAMES, Gareth et al. **An introduction to statistical learning**. New York: springer, 2013.

GANESH, Prakhar. Growing your own RNN cell : Simplified. *In: Growing your own RNN cell : Simplified*. [S. l.], 10 jun. 2019. Disponível em: <https://towardsdatascience.com/growing-your-own-rnn-cell-simplified-b68ba2c0f082>. Acesso em: 22 jan. 2020.

GREFF, Klaus et al. LSTM: A search space odyssey. **IEEE transactions on neural networks and learning systems**, v. 28, n. 10, p. 2222-2232, 2016.

GRUBER, Martin. Another Puzzle: The Growth in Actively Managed Mutual Funds. **The Journal of Finance**, v. 51, n. 3, p. 783-810, 1996.

GU, Shihao; KELLY, Bryan; XIU, Dacheng. **Empirical asset pricing via machine learning**. National Bureau of Economic Research, 2018.

HOU, Lu et al. Normalization helps training of quantized lstm. In: Advances in Neural Information Processing Systems. 2019. p. 7346-7356.

HAMILTON, William P. The stock market barometer: a study of its forecast value based on Charles H. **Dow's theory of the price movement**. **Barron's, New York**, 1922.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural computation**, v. 9, n. 8, p. 1735-1780, 1997.

HUANG, Paoyu; LIAO, Yi-Ching; NI, Yensen. Momentum in the Chinese stock market: Evidence from stochastic oscillator indicators. **Emerging Markets Finance and Trade**, v. 51, n. sup1, p. S99-S110, 2015.

KAASTRA, Ieabeling; BOYD, Milton. Designing a neural network for forecasting financial. **Neurocomputing**, v. 10, p. 215-236, 1996.

KIM, Kyoung-jae. Financial time series forecasting using support vector machines. **Neurocomputing**, v. 55, n. 1-2, p. 307-319, 2003.

KROLLNER, Bjoern; VANSTONE, Bruce J.; FINNIE, Gavin R. Financial time series forecasting with machine learning techniques: a survey. In: **Esann**. 2010.

JUNG, Jin-Kyu; PATNAM, Manasa; TER-MARTIROSYAN, Anna. **An algorithmic crystal ball: Forecasts-based on machine learning**. International Monetary Fund, 2018.

LAURENT, César et al. **Batch normalized recurrent neural networks**. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016. p. 2657-2661.

LO, Andrew W.; MAMAYSKY, H.; WANG, J. **Foundations of technical analysis**. Journal of Finance, ed. 55, p. 1705-1765, 2000.

LO, Andrew W.; HASANHODZIC, Jasmina. **The heretics of finance: Conversations with leading practitioners of technical analysis**. John Wiley and Sons, 2010.

LO, Andrew W. **Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis**. Journal of investment consulting, v. 7, n. 2, p. 21-44, 2005.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. Revista de Informática Teórica e Aplicada, vol.14, no2, pp 43-67, 2007. Disponível em: <http://seer.ufrgs.br/index.php/rita/article/viewPDFInterstitial/rita_v14_n2_p43-67/3543>. Acesso em: 16 out. 2010

MAJIDI, Babak; PATRA, Jagdish C; SHAMISA, Abdolah. Sliding-window-based real-time model order reduction for stability prediction in smart grid. **IEEE Transactions on Power Systems**, v. 34, n. 1, p. 326-337, 2018.

MARKOWITZ, Harry. Portfolio selection. **The Journal of Finance**, v. 7 n. 1, p. 77-91, 1952.

MARTINS, Maria Eugénia Graça. Desvio padrão amostral. **Revista de ciência elementar**, v. 1, n. 1, 2013.

MCKINNEY, Wes. **Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython**. [S. l.]: O'Reilly Media, Inc., 2012.

MCNALLY, Sean; ROCHE, Jason; CATON, Simon. Predicting the price of bitcoin using machine learning. In: **2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)**. IEEE, 2018. p. 339-343.

MENKHOFF, Lukas. The use of technical analysis by fund managers: International evidence. **Journal of Banking & Finance**, v. 34, n. 11, p. 2573-2586, 2010.

MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python: A Guide for Data Scientists**: A Guide for Data Scientists. [S. l.]: O'Reilly Media, Inc, 2016.

MURPHY, John J. **Technical analysis of the financial markets: A comprehensive guide to trading methods and applications**. Penguin, 1999.

NAZÁRIO, Rodolfo Toríbio Farias et al. **A literature review of technical analysis on stock markets**. The Quarterly Review of Economics and Finance, v. 66, p. 115-126, 2017.

NELSON, David MQ; PEREIRA, Adriano CM; DE OLIVEIRA, Renato A. **Stock market's price movement prediction with LSTM neural networks**. In: 2017 International joint conference on neural networks (IJCNN). IEEE, 2017. p. 1419-1426.

NEMATİ S, Holder *et al.* An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. **Critical Care Medicine**, [s. l.], v. 46, p. 547-553, 2018.

OBEIDAT, Samer et al. Adaptive portfolio asset allocation optimization with deep learning. **International Journal on Advances in Intelligent Systems**, v. 11, n. 1, p. 25-34, 2018.

PARK, Cheol-Ho; IRWIN, Scott H. What do we know about the profitability of technical analysis?. **Journal of Economic Surveys**, v. 21, n. 4, p. 786-826, 2007.

PERSON, John L. **Candlestick and Pivot Point Trading Triggers,+ Website: Setups for Stock, Forex, and Futures Markets**. John Wiley & Sons, 2006.

RAVICHANDRA, Thangjam; CHACKO, Ashish; GANGULY, Shruti. **TECHNICAL ANALYSIS OF EQUITY SHARES**. **International Journal of Research in Finance & Marketing**, p. 10-17, 2015.

JAIN, Abhijit; KONKU, Daniel; RAYOME, David L. Technical analysis: Donchian channels and the british pound. In: **IABE-2007 Annual Conference**. 2007. p. 302.
report, National Bureau of Economic Research.

ROSILLO, Rafael; DE LA FUENTE, David; BRUGOS, José A. Lopez. Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies. *Applied Economics*, 2013, 45.12: 1541-1550.

ROTELLA, R. P. (1992) *The Elements of Successful Trading*. Institute of Finance Simon & Schuster. New York:

SANTOS, Guilherme. **Uma aplicação de redes neurais recorrentes do tipo LSTM à previsão dos preços de curto prazo do mercado de energia elétrica brasileiro**. 2019. Tese de Doutorado.

SATCHWELL, Christopher J. **Technical analysis formation recognition using pivot points**. U.S. Patent n. 7,835,966, 16 nov. 2010.

SCHWAGER, Jack D. **Technical analysis**. John Wiley & Sons, 1995.

SENDHIL, Mullainathan; SPIESS, Jann. **Machine Learning: An Applied Econometric Approach**. *Journal of Economic Perspectives*, ed. 31, p. 87-106. 2017.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding machine learning: From theory to algorithms**. Cambridge university press, 2014.

SHARPE, William F. The Sharpe Ratio. **The Journal of Portfolio Management**, [s. l.], 1964.

SHIRYAEV, Albert; XU, Zuoquan; ZHOU, Xun Yu. Thou shalt buy and hold. **Quantitative finance**, v. 8, n. 8, p. 765-776, 2008.

STECKLER, David R. Detecting New Trends Early. **TECHNICAL ANALYSIS OF STOCKS AND COMMODITIES-MAGAZINE EDITION-**, v. 17, p. 60-67, 1999.

THOMSETT, Michael C. **CMF--Chaikin Money Flow: Changes Anticipating Price Reversal**. FT Press, 2010.

VARIAN, H. R. **Big data: New tricks for econometrics**. The Journal of Economic Perspectives, ed. 28, p. 3–27, 2014.

VUKOVIC, Darko et al. Neural network forecasting in prediction Sharpe ratio: Evidence from EU debt market. **Physica A: Statistical Mechanics and its Applications**, v. 542, p. 123331, 2020.

VYKLYUK, Yaroslav; VUKOVIC, Darko; JOVANOVIC, Ana. Forex prediction with neural network: USD/EUR currency pair. **Актуальні проблеми економіки**, n. 10, p. 261-273, 2013.

WILDER, J. Welles. **New concepts in technical trading systems**. Trend Research, 1978.

WORKING, Holbrook. **A random-difference series for use in the analysis of time series**. journal of the American Statistical Association, 1934, 29.185: 11-24.

YUNPENG, Liu et al. **Multi-step ahead time series forecasting for different data patterns based on LSTM recurrent neural network**. In: 2017 14th Web Information Systems and Applications Conference (WISA). IEEE, 2017. p. 305-310.

ZHU, Yingzi; ZHOU, Guofu. Technical analysis: An asset allocation perspective on the use of moving averages. **Journal of Financial Economics**, v. 92, n. 3, p. 519-544, 2009.

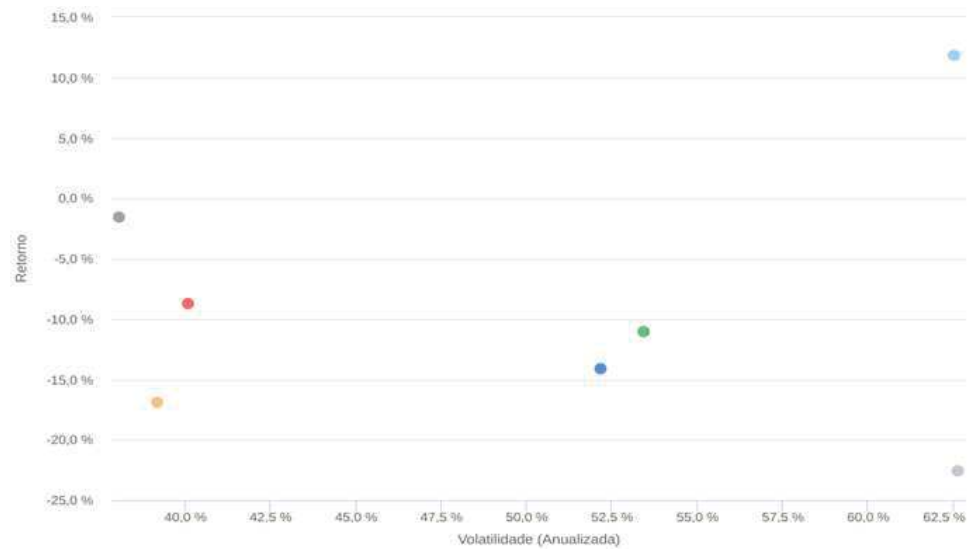
CAPÍTULO 6 ANEXO

Anexo 1 - Comparativo entre os retornos acumulados dos portfólios que efetuam a maximização do índice de Sharpe, dos portfólios que minimizam a volatilidade e do Ibovespa, entre as datas de 01/01/2020 a 31/08/2020.



Retorno acumulado - 01/01/2020 até 31/08/2020		
	Ativo	Retorno
	Max Sharpe 3d	11,88 %
	Min Vol 3d	-1,53 %
	Min Vol 4d	-8,70 %
	Max Sharpe 4d	-11,08 %
	Ibovespa	-14,07 %
	Min Vol 5d	-16,86 %
	Max Sharpe 5d	-22,59 %

Anexo 2 - Relação risco versus retorno dos portfólios que minimizam a volatilidade, portfólios que maximizam o índice de Sharpe e do Ibovespa, entre as datas de 01/01/2020 a 31/08/2020.



Risco X Retorno de 01/01/2020 até 31/08/2020			
Ativo	Retorno	Volatilidade (Anualizada)	Índice de Sharpe
Max Sharpe 3d	11,88 %	62,52 %	0,11
Min Vol 3d	-1,53 %	38,08 %	-0,17
Ibovespa	-14,07 %	52,16 %	-0,27
Max Sharpe 4d	-11,08 %	53,44 %	-0,30
Min Vol 4d	-8,70 %	40,10 %	-0,34
Max Sharpe 5d	-22,59 %	62,63 %	-0,44
Min Vol 5d	-16,86 %	39,19 %	-0,56

Anexo 3 - Gráfico Drawdown (Perdas). Mínima Volatilidade x Máximo Sharpe x Ibovespa (01/01/2019 – 31/08/2020).



Draw Down - 01/01/2019 à 31/08/2020				
Ativo	Início do Máximo Draw Down	Data do Máximo Draw Down	Máximo Draw Down	
Max Sharpe 5d	23/01/2020	18/03/2020	-58,84 %	
Max Sharpe 3d	23/01/2020	18/03/2020	-50,98 %	
Max Sharpe 4d	22/01/2020	18/03/2020	-48,13 %	
Ibovespa	23/01/2020	23/03/2020	-46,82 %	
Min Vol 5d	22/01/2020	23/03/2020	-41,54 %	
Min Vol 4d	22/01/2020	23/03/2020	-39,82 %	
Min Vol 3d	22/01/2020	23/03/2020	-33,07 %	

Anexo 4 - Tabela com rentabilidades mensais. Mínima Volatilidade x Máximo Sharpe x Ibovespa (01/2019 – 08/2020).

Mês, Ano	Ibovespa	Min Vol 3d	Min Vol 4d	Min Vol 5d	Max Sharpe 3d	Max Sharpe 4d	Max Sharpe 5d
janeiro, 2019	10,8 %	9,0 %	9,0 %	8,1 %	14,0 %	19,2 %	13,4 %
fevereiro, 2019	-1,9 %	-3,0 %	-2,2 %	-1,4 %	-5,6 %	-5,9 %	-1,9 %
março, 2019	-0,2 %	-0,6 %	0,0 %	0,3 %	3,3 %	3,8 %	1,9 %
abril, 2019	1,0 %	1,2 %	2,6 %	2,3 %	11,6 %	10,0 %	9,7 %
maio, 2019	0,7 %	-0,6 %	0,8 %	-0,6 %	2,3 %	6,7 %	3,4 %
junho, 2019	4,1 %	3,6 %	3,4 %	3,7 %	5,0 %	2,5 %	2,4 %
julho, 2019	0,8 %	3,5 %	2,8 %	3,4 %	15,4 %	11,2 %	9,8 %
agosto, 2019	-0,7 %	1,2 %	2,5 %	1,0 %	7,3 %	8,8 %	4,9 %
setembro, 2019	3,6 %	4,6 %	2,5 %	3,0 %	2,9 %	4,2 %	4,5 %
outubro, 2019	2,4 %	-3,0 %	-2,9 %	-3,6 %	2,6 %	-4,5 %	1,0 %
novembro, 2019	0,9 %	7,5 %	5,9 %	5,4 %	3,2 %	2,9 %	3,6 %
dezembro, 2019	6,8 %	5,8 %	4,8 %	6,1 %	3,4 %	3,7 %	5,5 %
janeiro, 2020	-1,6 %	2,1 %	5,3 %	2,9 %	5,4 %	4,7 %	6,0 %
fevereiro, 2020	-8,4 %	-7,9 %	-11,2 %	-12,5 %	-9,6 %	-12,3 %	-12,3 %
março, 2020	-29,9 %	-16,9 %	-23,7 %	-25,0 %	-30,3 %	-27,2 %	-42,3 %
abril, 2020	10,3 %	5,0 %	4,4 %	3,6 %	20,1 %	15,2 %	18,7 %
maio, 2020	8,6 %	6,0 %	7,1 %	6,3 %	20,7 %	6,1 %	7,7 %
junho, 2020	8,8 %	3,0 %	3,9 %	5,0 %	8,2 %	5,2 %	13,3 %
julho, 2020	8,3 %	10,0 %	7,8 %	9,1 %	9,5 %	9,9 %	5,7 %
agosto, 2020	-3,4 %	0,3 %	0,9 %	-4,1 %	4,2 %	-1,0 %	-5,7 %

Anexo 5 - Composição portfólio Mínimo Volatilidade, com previsões de cinco dias

Empresa	Código	Part. (%)
AMBEV S/A	ABEV3	10,31
BBSEGURIDADE	BBSE3	4,88
BRF SA	BRFS3	3,46
CIELO	CIEL3	0,32
CEMIG	CMIG4	1,71
COGNA ON	COGN3	1,54
GERDAU	GGBR4	2,07
IRBBRASIL RE	IRBR3	9,60
P.ACUCAR-CBD	PCAR4	3,12
LOCALIZA	RENT3	10,08
SUZANO S.A.	SUZB3	9,52
ULTRAPAR	UGPA3	5,10
VALE	VALE3	8,00
TELEF BRASIL	VIVT4	30,29

Fonte: Autor

Anexo 6 - Composição portfólio Máximo Sharpe, com previsões para cinco dias

Empresa	Código	Part. (%)
IRBBRASIL RE	IRBR3	20,09
COSAN	CSAN3	19,04
BRF SA	BRFS3	17,73
MAGAZ LUIZA	MGLU3	12,83
COGNA ON	COGN3	9,93
RUMO S.A.	RAIL3	7,35
B3	B3SA3	5,04
NATURA	NATU3	3,68
TELEF BRASIL	VIVT4	1,96
SID NACIONAL	CSNA3	1,21
CCR SA	CCRO3	1,14

Fonte: Autor

Anexo 7 - Composição portfólio Mínimo Volatilidade, com previsões para quatro dias

Empresa	Código	Part. (%)
AMBEV S/A	ABEV3	3,50
AZUL	AZUL4	0,59
PETROBRAS BR	BRDT3	0,26
BRF SA	BRFS3	3,02
CEMIG	CMIG4	7,33
GERDAU	GGBR4	0,79
GERDAU MET	GOAU4	3,81
IRBBRASIL RE	IRBR3	6,60
ITAUUNIBANCO	ITUB4	4,22
JBS	JBSS3	2,98
KLABIN S/A	KLBN11	16,87
LOCALIZA	RENT3	7,25
SANTANDER BR	SANB11	1,11
SUZANO S.A.	SUZB3	3,85
ULTRAPAR	UGPA3	2,75
VALE	VALE3	1,88
TELEF BRASIL	VIVT4	33,18

Fonte: Autor

Anexo 8 - Composição portfólio Máximo Sharpe, com previsões para quatro dias

Empresa	Código	Part. (%)
AZUL	AZUL4	1,13
B3	B3SA3	0,14
BRF SA	BRFS3	9,71
CCR SA	CCRO3	5,55
COSAN	CSAN3	9,67
SID NACIONAL	CSNA3	0,44
ELETRONBRAS	ELET6	4,51
IRBBRASIL RE	IRBR3	10,18
JBS	JBSS3	19,69
MAGAZ LUIZA	MGLU3	4,34
MRV	MRVE3	9,04
RUMO S.A.	RAIL3	9,75
TELEF BRASIL	VIVT4	11,48
VIAVAREJO	VVAR3	2,75
WEG	WEGE3	1,61

Fonte: Autor

Anexo 9 - Composição portfólio Mínimo Volatilidade, com previsões para três dias.

Empresa	Código	Part. (%)
BBSEGURIDADE	BBSE3	7,41
PETROBRAS BR	BRDT3	3,30
BRF SA	BRFS3	4,92
ITAUUNIBANCO	ITUB4	1,51
KLABIN S/A	KLBN11	6,05
NATURA	NATU3	6,28
P.ACUCAR-CBD	PCAR4	1,23
RUMO S.A.	RAIL3	5,87
SUZANO S.A.	SUZB3	13,68
ULTRAPAR	UGPA3	5,22
VALE	VALE3	5,11
TELEF BRASIL	VIVT4	37,05
WEG	WEGE3	2,37

Fonte: Autor

Anexo 10 - Composição portfólio Máximo Sharpe, com previsões para três dias

Empresa	Código	Part. (%)
AZUL	AZUL4	2,00
B3	B3SA3	6,65
BRF SA	BRFS3	16,62
CCR SA	CCRO3	2,38
SID NACIONAL	CSNA3	1,19
IRBBRASIL RE	IRBR3	6,59
JBS	JBSS3	7,62
MAGAZ LUIZA	MGLU3	24,83
MRV	MRVE3	1,20
RUMO S.A.	RAIL3	14,69
SABESP	SBSP3	10,54
VIAVAREJO	VVAR3	4,76
WEG	WEGE3	0,93

Fonte: Autor