



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Londrina



RAÍSA NATALIA FRANÇA LINS

**PREVISÃO DE FUNDOS DE INVESTIMENTOS COM O USO DE MACHINE
LEARNING**

Londrina

2020

RAÍSA NATALIA FRANÇA LINS

**PREVISÃO DE FUNDOS DE INVESTIMENTOS COM O USO DE MACHINE
LEARNING**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Engenharia de Produção, do Departamento Acadêmico de Engenharia de Produção da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Bruno Samways dos Santos

Londrina

2020

TERMO DE APROVAÇÃO

TRABALHO DE CONCLUSÃO DE CURSO - TCC PREVISÃO DE FUNDOS DE INVESTIMENTOS COM O USO DE MACHINE LEARNING

Por
Raísa Natalia França Lins

Monografia apresentada às 9 horas e 0 min. do dia 30 de novembro de 2020 como requisito parcial, para conclusão do Curso de Bacharelado em Engenharia de Produção da Universidade Tecnológica Federal do Paraná, Campus Londrina. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação e conferidas, bem como achadas conforme, as alterações indicadas pela Banca Examinadora, o trabalho de conclusão de curso foi considerado APROVADO.

Banca examinadora:

Prof. Dr. Marcos J. G. Rambalducci	Membro
Prof. Dr. Rogério Tondato	Membro
Prof. Dr. Bruno Samways dos Santos	Orientador
Profa. Dra. Silvana Rodrigues Quintilhano	Professor(a) responsável TCCII



Documento assinado eletronicamente por (Document electronically signed by) **BRUNO SAMWAYS DOS SANTOS, PROFESSOR DO MAGISTERIO SUPERIOR**, em (at) 08/12/2020, às 13:35, conforme horário oficial de Brasília (according to official Brasilia-Brazil time), com fundamento no (with legal based on) art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por (Document electronically signed by) **MARCOS JERONIMO GOROSKI RAMBALDUCCI, PROFESSOR DO MAGISTERIO SUPERIOR**, em (at) 08/12/2020, às 16:15, conforme horário oficial de Brasília (according to official Brasilia-Brazil time), com fundamento no (with legal based on) art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por (Document electronically signed by) **ROGERIO TONDATO, PROFESSOR DO MAGISTERIO SUPERIOR**, em (at) 08/12/2020, às 20:56, conforme horário oficial de Brasília (according to official Brasilia-Brazil time), com fundamento no (with legal based on) art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por (Document electronically signed by) **SILVANA RODRIGUES QUINTILHANO TONDATO, PROFESSOR DO MAGISTERIO SUPERIOR**, em (at) 10/12/2020, às 10:56, conforme horário oficial de Brasília (according to official Brasilia-Brazil time), com fundamento no (with legal based on) art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site (The authenticity of this document can be checked on the website) https://sei.utfpr.edu.br/sei/controlador_externo.php?cao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador (informing the verification code) 1774252 e o código CRC (and the CRC code) A3D0449CB.

AGRADECIMENTOS

Agradeço, primeiramente, à Deus pela oportunidade de ter trilhado essa jornada e por todo o cuidado junto a mim e à minha família, principalmente ao meu pai e minha mãe, que possibilitaram essa experiência única, sempre me encorajando a seguir e a batalhar pelos meus sonhos.

Agradeço aos meus amigos que fizeram parte dessa caminhada, dividindo os bons e maus momentos, criando memórias que ficaram eternizadas em minha mente e coração.

E também ao professor Bruno Samways que, principalmente nesse ano difícil de pandemia, me ajudou a me manter firme e me auxiliou em todos os aspectos para a execução desse trabalho.

LINS, Raísa Natalia França. Previsão de Fundos de Investimentos com o uso de Machine Learning. 74 p. Trabalho de Conclusão de Curso. Graduação em Engenharia de Produção. Universidade Tecnológica Federal do Paraná

RESUMO

O mercado financeiro é amplo e entrega ao investidor diversas opções de investimentos, como os produtos de renda fixa, ações e fundos de investimentos, por exemplo. Para o caso dos brasileiros, leva-se em consideração a segurança de sua aplicação e sua rentabilidade. Porém, para os fundos de investimentos é possível, estimar qual será a rentabilidade futura da aplicação, tendo apenas os dados históricos como uma referência de comportamento. Neste contexto, o presente trabalho teve como objetivo analisar técnicas de *Machine Learning* para estimar variações futuras de cotas em três características de fundos: conservador, moderado e agressivo. Para tal, foram utilizadas Rede Neural e Regressão Linear, com diferentes conjuntos de treinamento, ou seja, foram separados um total de nove conjunto de dados de séries históricas de variação de cota e analisados por três métricas diferentes: erro absoluto médio, erro quadrático médio e raiz do erro quadrático médio. A partir da ferramenta Weka (*Waikato Environment for Knowledge Analysis*), com a extensão *Forecast*, os algoritmos foram treinados e testados, destaca-se a técnica de Redes Neurais a qual teve um melhor desempenho para os diferentes fundos e conjuntos de treinamento. Entretanto, a Regressão Linear obteve resultados significativos para a previsão da variação da cota com os conjuntos de treinamento com um ano de dados de séries históricas.

Palavras-chave: Fundo de Investimento. Previsão. Machine Learning. Rede Neural. Regressão Linear.

LINS, Raísa Natalia França. Forecasting Investment Funds using Machine Learning. 74 p. Undergraduate Final Project. Industrial Engineering Program. Federal University of Technology – Paraná.

ABSTRACT

The financial market is broad and provides to the investor with several investment options, such as fixed income products, stocks and investment funds, for example. For the case of Brazilians, the security of its application and its profitability are taken into account. However, for investment funds it is possible to estimate the future profitability of the application, using only historical data as a reference for its behavior. In this context, the present study aimed to analyze the Machine Learning techniques to forecast future variations of shares in three fund types: conservative, moderate and aggressive. For this purpose, Neural Network and Linear Regression were used, with different training sets, i.e., a total of nine data sets of historical series of quota and destination variation were separated by three different metrics: mean absolute error, mean square error and root of the mean square error. Using the software Weka (Waikato Environment for Knowledge Analysis), with the Forecast extension, the algorithms were trained and tested, the Neural Networks technique stands out, which had a better performance for the different funds and training sets. However, the Linear Regression obtained interesting results when forecasting the quota variation with the training sets using one year of historical data.

Keywords: *Investment Fund. Forecast. Machine Learning. Neural Network. Linear Regression.*

LISTA DE FIGURAS

Figura 1 - Métodos e ferramentas da mineração de dados	30
Figura 2 - Processo de descoberta de conhecimento em base de dados.....	31
Figura 3 - Representação biológica do neurônio.....	34
Figura 4 - Neurônio artificial.	35
Figura 5 - Representação da entrada de um neurônio artificial.....	36
Figura 6 - Função de ativação Sigmoidal	36
Figura 7 - RNA com múltiplas camadas	37
Figura 8 - Passo a passo da pesquisa	47
Figura 9 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo Conservador.	53
Figura 10 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo Conservador	54
Figura 11 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de completa do fundo Conservador.....	55
Figura 12 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo Moderado.....	55
Figura 13 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo Moderado	56
Figura 14 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de completa do fundo Moderado	57
Figura 15 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo Agressivo	58
Figura 16 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo Agressivo.....	59
Figura 17 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de completa do fundo Agressivo	59
Figura 18 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Conservador	60
Figura 19 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Conservador.....	61
Figura 20 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Conservador	62

Figura 21 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Moderado	63
Figura 22 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Moderado	64
Figura 23 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Moderado.....	64
Figura 24 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Agressivo.....	65
Figura 25 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Agressivo	66
Figura 26 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Agressivo	67
Figura 27 - Heatmap do erro MAE para os algoritmos de previsão.....	67
Figura 28 - Heatmap do erro RMSE para os algoritmos de previsão	67
Figura 29 - Heatmap do erro MSE para os algoritmos de previsão.....	68

LISTA DE QUADROS

Quadro 1 - Outros termos em Fundos de Investimentos.....	26
Quadro 2 - Trabalhos correlatos.....	42

LISTA DE TABELAS

Tabela 1 - Teste da Taxa de Aprendizagem 0,1 para o fundo Conservador.....	49
Tabela 2 - Teste da Taxa de Aprendizagem 0,2 para o fundo Conservador.....	50
Tabela 3 - Teste da Taxa de Aprendizagem 0,3 para o fundo Conservador.....	50
Tabela 4 - Teste da Taxa de Aprendizagem 0,1 para o fundo Moderado.....	50
Tabela 5 - Teste da Taxa de Aprendizagem 0,2 para o fundo Moderado.....	51
Tabela 6 - Teste da Taxa de Aprendizagem 0,3 para o fundo Moderado.....	51
Tabela 7 - Teste da Taxa de Aprendizagem 0,1 para o fundo Agressivo.....	52
Tabela 8 - Teste da Taxa de Aprendizagem 0,2 para o fundo Agressivo.....	52
Tabela 9 - Teste da Taxa de Aprendizagem 0,3 para o fundo Agressivo.....	52
Tabela 10 - Algoritmo de Regressão Linear aplicado a base de um ano do fundo Conservador.....	53
Tabela 11 - Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Conservador.....	54
Tabela 12 - Algoritmo de Regressão Linear aplicado a base completa do fundo Conservador.....	54
Tabela 13 - Algoritmo de Regressão Linear aplicado a base de um ano do fundo Moderado.....	55
Tabela 14 - Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Moderado.....	56
Tabela 15 - Algoritmo de Regressão Linear aplicado a base completa do fundo Moderado.....	57
Tabela 16 - Algoritmo de Regressão Linear aplicado a base de um ano do fundo Agressivo.....	57
Tabela 17 - Outros termos em Fundos de Investimentos Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Agressivo.	58
Tabela 18 - Algoritmo de Regressão Linear aplicado a base completa do fundo Agressivo.....	59
Tabela 19 - Algoritmo de Rede Neural aplicado a base de um ano do fundo Conservador.....	60
Tabela 20 - Algoritmo de Rede Neural aplicado a base de dois anos do fundo Conservador.....	61
Tabela 21 - Algoritmo de Rede Neural aplicado a base completa do fundo Conservador.....	62

Tabela 22 - Algoritmo de Rede Neural aplicado a base de um ano do fundo Moderado.	62
Tabela 23 - Algoritmo de Rede Neural aplicado a base de dois anos do fundo Moderado.	63
Tabela 24 - Algoritmo de Rede Neural aplicado a base completa do fundo Moderado	64
Tabela 25 - Algoritmo de Rede Neural aplicado a base de um ano do fundo Agressivo	65
Tabela 26 - Algoritmo de Rede Neural aplicado a base de dois anos do fundo Agressivo.	66
Tabela 27 - Algoritmo de Rede Neural aplicado a base completa do fundo Agressivo.	66

LISTA DE ABREVIATURAS E SIGLAS

CDC	Center for Disease Control
AMBIMA	Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais
ARFF	<i>Attribute Relation File Format</i>
ARIMA	<i>Autoregressive Integrated Moving Average</i>
BACEN	Banco Central
BDRS	<i>Brazilian Depositary Receipts</i>
BRL	Real Brasileiro
B3	Brasil, Bolsa e Balcão
CDB	Certificado de Depósitos Bancários
CDI	Certificado de Depósitos Interfinanceiros
CETES	Rendimentos sobre Certificados do Tesouro
CMN	Conselho Monetário Nacional
CNPC	Conselho Nacional de Previdência Complementar
CNSP	Conselho Nacional de Seguros Privados
CPI	Índice nacional de preços
CTVM	Corretoras de Títulos e Valores Mobiliários
CVM	Comissão de Valores Mobiliários
DJIA	Índice <i>Dow Jones Industrial Average</i>
DVTM	Distribuidoras de Títulos e Valores Mobiliários
FGC	Fundo Garantidor de Crédito
FIDC	Fundo de Investimento em Direito Creditório
FRPCA	<i>Fast Robust Principal Component Analysis</i>
GA	Algoritmo Genético
GARCH	Redes Neurais e Arima
GNU	General Public License
PHM	<i>Prognostic and Health Monitoring</i>
IBOV	Índice Bovespa
IGP-DI	Índice Geral de Preços - Disponibilidade Interna
IOF	Imposto sobre Operações Financeiras
IPCA	Índice Nacional de Preços ao Consumidor Amplo
IR	Imposto de Renda
JCP	Juros sobre Capital Próprio

JGP	<i>Jakurski, Guedes & Partners</i>
KDD	<i>Knowlegde Discovery in Databases</i>
KPCA	<i>Kernel Principal Component Analysis</i>
LCA	Letra de Crédito de Agronegócio
LCI	Letra de Crédito Imobiliário
LFT	Letra Financeira do Tesouro
LTN	Letra do Tesouro Nacional
M1	Agregado Monetário
MAE	Erro Absoluto Médio
MAPE	Erro Absoluto Médio Percentual
MSE	Erro Quadrático Médio
NAV	Valor Patrimonial Líquido
NIPC	Índice Nacional de Preços ao Consumidor
NTN-B	Nota do Tesouro Nacional – Série B
NTN-B Principal	Nota do Tesouro Nacional - Série B Principal
NTN-F	Notas do Tesouro Nacional – Série F
PCA	<i>Kernel Principal Component Analysis</i>
PIB	Produto Interno Bruto
PREVIC	Superintendência de Seguros Privados e pela Superintendência Nacional de Previdência Complementar
RI	Reservas Internacionais
RL	Régressão Linear
RMSE	Raiz do Erro Quadrático Médio
RNA	Redes Neurais Artificiais
SDRM	Risco de Default Soberano do México
SFN	Sistema Financeiro Nacional
SIM	Modelo de Índice Único
TIC	Coeficiente de Desigualdade de Theil
USD MXN	Dólar Americano Peso Mexicano
USD	Dólar Americano

SUMÁRIO

1. INTRODUÇÃO.....	12
1.1. Objetivo geral.....	13
1.2. Objetivos específicos.....	13
1.3. Justificativa.....	13
1.4. Metodologia de Pesquisa.....	14
1.5. Estrutura do trabalho.....	14
2. REFERENCIAL TEÓRICO.....	16
2.1. Investimentos.....	16
2.1.2. Tipos de produtos.....	20
2.2. Fundos de Investimentos.....	24
2.2.1. Tipos de Fundos de Investimentos.....	27
2.3. Mineração de Dados.....	29
2.3.1. Séries Temporais.....	32
2.3.2. Regressão linear.....	33
2.3.3. Redes neurais artificiais.....	33
2.3.4. Métrica de avaliação.....	38
2.4. Trabalhos correlatos.....	39
3. METODOLOGIA.....	42
3.1 Descrição do conjunto de dados.....	43
3.2 Sequência da pesquisa.....	45
3.3 Ferramentas utilizadas.....	48
4. RESULTADOS E DISCUSSÕES.....	49
4.1 Teste de Parâmetros.....	49
4.1 Regressão Linear.....	53
4.2 Rede Neural.....	60
5. CONCLUSÕES.....	69
REFERÊNCIAS BIBLIOGRÁFICAS.....	71

1. INTRODUÇÃO

A pesquisa intitulada Raio X do investidor brasileiro, realizada pela Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (AMBIMA), com auxílio do Datafolha, divulgada em julho de 2018, mostrou qual é a realidade do brasileiro em relação aos investimentos. Inicialmente, mostrou que apenas 32% da população conseguiu economizar no ano de 2017 e, nesta ação de economizar, apenas 42% das pessoas entrevistadas aplicaram em produtos financeiros. A outra parte utilizou o dinheiro para adquirir bens como carro e apartamento, ou investiu em estudos e em viagens. Esse resultado demonstra que o brasileiro apresenta como visão de investimento um conceito mais amplo, que vai além de produtos financeiros.

Essa pesquisa também comprovou que a poupança é o investimento preferido dos investidores, com 89% de adeptos. Como é simples e de fácil acesso, se tornou o principal investimento do brasileiro, porém, ao considerar o fator histórico, com alta taxa de inflação e mínima histórica da taxa básica de juros, o rendimento real da poupança acaba estando próximo de zero ou negativo. Já a outra parte dos entrevistados se dividem, pois 5% investe em fundos de investimentos, 4% em títulos privados e 3% títulos públicos, os quais são investimentos financeiros que mostram historicamente um ganho real.

Além disso, tal pesquisa mostrou os motivos e fatores que levam o investidor a decidir o tipo de investimento, em que a maioria investe considerando a segurança, com 54% dos entrevistados. Já os que consideram a rentabilidade dos investimentos aparecem com 16% e 12% consideram o fator liquidez, que é a possibilidade de resgatar o dinheiro em um pequeno intervalo de tempo.

Saindo da ótica de onde a população brasileira investe e o motivo que leva a escolher tal investimento, surge o cenário de que há diversos tipos de produtos de investimentos, no qual existem taxas de rentabilidade variadas, possibilidade de liquidez ou não, vencimento, entre outras, deixando o investidor ainda mais confuso em onde deve aplicar seus recursos. Neste cenário de variados tipos de produtos, há a indústria de fundos brasileira, que conta atualmente com cerca de 595 gestoras, segundo *ranking* divulgado pela AMBIMA (2019), envolvendo cerca de R\$ 3 trilhões, considerando apenas os fundos de renda fixa, ações, multimercados e cambial.

Os investimentos em fundos de investimentos possuem vantagem em relação a outras aplicações financeiras, pois o tempo de aplicação pode ser menor, o risco pode ser diluído, entre outros motivos. Porém, apesar de alguns fundos de investimentos apresentarem características constantes e conservadoras, não é possível afirmar qual será o rendimento de tal fundo.

Por fim, como se pode ver pela pesquisa divulgada pela AMBIMA, o investidor leva em consideração a rentabilidade e segurança de tal aplicação. E o fato de conseguir prever antecipadamente, por meio da mineração de dados, a rentabilidade dos fundos de investimentos, faz com que o retorno alcançado possa ser maior e evitar retiradas em momentos errados, por exemplo. Neste contexto, levanta-se o seguinte problema: Qual a capacidade de técnicas de *Machine Learning* prever os resultados de diferentes tipos de fundos de investimentos?

1.1. Objetivo geral

Analizar a capacidade de previsão da rentabilidade futura, a partir da variação da cota utilizando técnicas de *Machine Learning* em três fundos de investimentos de diferentes características (conservador, moderado e agressivo).

1.2. Objetivos específicos

O objetivo geral foi analisado a partir dos seguintes objetivos específicos:

- Investigar as aplicações de *Machine Learning* na área financeira;
- Construir modelos de *Machine Learning* para a previsão da rentabilidade de fundos de investimentos;
- Comparar os resultados obtidos a partir dos erros de previsão gerados.

1.3. Justificativa

Conforme Solin et al. (2019), no mercado financeiro, os investidores enfrentam vários obstáculos, como os retornos incertos e altos riscos no investimento. Considerando este cenário do mercado financeiro e a realidade dos brasileiros, conforme o tipo de investimento que investe e o motivo dessas escolhas, é

interessante buscar uma melhor rentabilidade e um menor risco, porém isso só é possível por meio de previsões da rentabilidade futura desse investimento. Sendo assim, tendo ferramentas que trabalhem com a previsão de rentabilidade, pode-se ter maiores ganhos, aliado a um maior conforto e segurança.

1.4. Classificação da Pesquisa

Em relação à natureza desta pesquisa, trata-se de um estudo quantitativo. Para Gerhardt e Silveira (2009), a pesquisa quantitativa tem como intuito o raciocínio dedutivo, as regras da lógica e os atributos mensuráveis da experiência humana.

Quanto aos objetivos, o estudo em questão é sobre a pesquisa exploratória, no qual proporciona maiores informações sobre o assunto que será investigado. Segundo Gil (2008), esse tipo de pesquisa tem como finalidade desenvolver, esclarecer e modificar conceitos e ideias, considerando a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores.

Em relação ao método de pesquisa, trata-se de modelagem e simulação. Modelagem é uso de técnicas para elaborar um modelo matemático ou descrever o funcionamento de um sistema produtivo ou parte dele e simulação é o uso de técnicas matemáticas e computacionais para simular o funcionamento de sistemas produtivos (BERTO *et al.*, 2014).

1.5. Estrutura do trabalho

Após o capítulo introdutório, em que foi apresentada contextualização sobre o assunto, os objetivos gerais e específicos, a importância desse trabalho e a metodologia de pesquisa utilizada, é apresentado o referencial teórico.

Inicialmente aborda-se o tema sobre investimentos, por meio de explicações acerca da estrutura do Sistema Financeiro Nacional, da forma de funcionamento dos investimentos e dos produtos oferecidos e suas características. Além disso, são apresentadas as características dos fundos de investimentos, a dinâmica, tipos e vantagem em se investir em um fundo de investimento, no qual teve papel principal na aplicação deste trabalho.

Ainda no referencial teórico, o terceiro tema desenvolvido foi a mineração de dados, além dos algoritmos de Regressão Linear e Rede Neural. Na mineração também são dissertadas as métricas de avaliação comumente utilizadas para previsões. E para finalizar, são apresentados os trabalhos correlatos, por meio de artigos que envolvem o mesmo assunto e serviram como base para a criação dessa monografia.

Em metodologia é descrito o conjunto de dados utilizado, o passo a passo do trabalho e um fluxograma representativo, e é finalizado com as ferramentas utilizadas. Os resultados obtidos através da aplicação dos métodos estudados são apresentados no quarto capítulo, além de apresentar análise e discussão desses resultados. Para finalizar, no quinto capítulo são expostas as conclusões e considerações sobre toda a pesquisa realizada, a qual demonstra sugestões e ideias do que pode vir a ser pesquisado em trabalhos futuros.

2. REFERENCIAL TEÓRICO

A presente seção descreve brevemente a estrutura responsável pelas demandas sobre investimentos no Brasil e sobre os tipos de produtos disponibilizados no mercado financeiro.

2.1. Investimentos

O Sistema Financeiro Nacional é um conjunto de instituições e instrumentos financeiros que tem como objetivo transferir os recursos dos agentes econômicos superavitários para os deficitários, considerando-se pessoas, empresas ou o governo, como agentes. Esse sistema é constituído por três grandes níveis: os órgãos normativos, os supervisores e os operadores:

O Sistema Financeiro Nacional (SFN) é formado por um conjunto de entidades e instituições que promovem a intermediação financeira, isto é, o encontro entre credores e tomadores de recursos. É por meio do sistema financeiro que as pessoas, as empresas e o governo circulam a maior parte dos seus ativos, pagam suas dívidas e realizam seus investimentos (BACEN, 2019, sem página)

Esse sistema é constituído por três grandes níveis: os órgãos normativos, os supervisores e os operadores:

Os órgãos normativos determinam regras gerais para o bom funcionamento do sistema. As entidades supervisoras trabalham para que os integrantes do sistema financeiro sigam as regras definidas pelos órgãos normativos. Os operadores são as instituições que oferecem serviços financeiros, no papel de intermediários (BACEN, 2019, sem página).

No nível normativo, existe o Conselho Monetário Nacional (CMN), responsável por criar as normas que tange a todo o mercado, além de estabelecer regras para o bom funcionamento do Sistema Financeiro Nacional como um todo e, segundo definição do Banco Central do Brasil (2019), possui como responsabilidade manter a estabilidade da moeda e o desenvolvimento econômico e social do país. Junto ao Conselho Monetário Nacional, há o Conselho Nacional de Seguros Privados (CNSP) e o Conselho Nacional de Previdência Complementar (CNPC), com o objetivo de criar as normas tangentes ao mercado de previdência privada e o de previdência complementar aberta.

O nível da supervisão é composto pelo Banco Central (BACEN), Comissão de Valores Mobiliários (CVM), Superintendência de Seguros Privados e pela

Superintendência Nacional de Previdência Complementar (PREVIC), responsáveis em executar as normas que são estabelecidas pelos órgãos normativos. O BACEN e a CVM são os principais órgãos desse nível, no qual o Banco Central, segundo sua própria definição, atua para regular a quantidade de moeda na economia, permitindo a estabilidade de preços e financeira do país, regulando e supervisionando as instituições financeiras. O CVM também deve fiscalizar, normatizar, disciplinar e desenvolver o mercado de valores mobiliários no Brasil.

Por fim, o nível operacional, conforme Banco Central (2019), é constituído pelos operadores: os Bancos e Caixas Econômicas, cooperativas de crédito, instituições de pagamentos, administradores de consórcios, corretoras e distribuidoras, outras instituições não bancárias, bolsa de valores, bolsa de mercadorias e futuros, seguradoras e resseguradoras, entidades abertas de previdência, sociedades de capitalização e entidades fechadas de previdência complementar (fundos de pensão).

Além dos níveis explorados anteriormente, no Sistema Financeiro Nacional existe o mercado financeiro, que é onde os investimentos financeiros são realizados e tem grande importância em relação ao processo de desenvolvimento econômico do país. Segundo Assaf Neto (2012), quatro mercados compõe o Mercado Financeiro: o Mercado Monetário, o Mercado de Crédito, o Mercado Cambial e o Mercado de Capitais.

O Mercado Monetário se refere às operações de prazo curto e curtíssimo, pois possibilita controle rápido e ágil da liquidez da economia e das taxas de juros. Entre os papéis negociados nesse mercado estão os Certificado de Depósitos Interfinanceiros (CDI), os Certificado de Depósitos Bancários (CDB), Caderneta de Poupança e os Títulos Públicos emitidos pelo governo federal, estadual e municipal.

O Mercado de Crédito se refere às operações de financiamento de ativos permanentes e capital de giro das empresas de curto e médio prazo. O Mercado Cambial envolve as operações de conversão de moeda e de um país para outro. E o Mercado de Capitais se refere às operações financeiras de médio e longo prazo, de prazo indeterminado, como, por exemplo, as operações em ações.

Segundo Lemes Junior *et al.* (2011), o mercado financeiro, que possibilita a captação e aplicação de recursos a curto e longo prazo, é operacionalizado por diferentes instituições financeiras, de natureza privada ou pública, as quais têm suas operações supervisionadas por órgão regulamentadores e fiscalizadores.

Nesse contexto, os intermediários financeiros, segundo Pinheiro (2002), têm como função a aproximação de agentes superavitários que possuem excesso de recursos, mas não têm oportunidade de investi-los em atividades produtivas, e os agentes deficitários que estão em situação inversa, ou seja, precisando desse recurso.

Como exemplo, existem as Corretoras de Títulos e Valores Mobiliários (CTVM) e as Distribuidoras de Títulos e Valores Mobiliários (DTVM), as quais atuam intermediando a negociação de títulos e valores mobiliários entre investidores e tomadores de recursos.

As corretoras e distribuidoras, na atividade de intermediação, oferecem serviços como plataformas de investimento pela internet (*home broker*), consultoria financeira, clubes de investimentos, financiamento para compra de ações (conta margem) e administração e custódia de títulos e valores mobiliários dos clientes. Na remuneração pelos serviços, essas instituições podem cobrar comissões e taxas (BACEN, 2019, sem página).

Acerca do mercado financeiro e dos produtos de investimentos, existem dois principais segmentos: Renda Fixa e Renda Variável. O primeiro é um investimento em que é possível dimensionar a remuneração no início da aplicação, sendo o valor nominal ou pela variação de uma taxa. Já na Renda Variável, a remuneração é incerta e não se pode saber a rentabilidade no início da operação. Considerando os investimentos em renda fixa, podem ser diferenciados em relação ao tipo de emissor, forma do cálculo da rentabilidade, imposto de renda e o Imposto sobre Operações Financeiras (IOF), prazo de vencimento e valor mínimo para o investimento.

Sobre os prazos de vencimento, os títulos de Renda Fixa são sempre emitidos com uma data de vencimento já estipulada, que será quando a instituição emissora devolverá ao investidor o valor principal, ou seja, o valor investido e mais os juros que foram acordados no início da operação. Esta data está ligada diretamente aos juros propostos para a operação, que geralmente, quanto maior um prazo de vencimento, maior é a taxa de juros.

Ainda em relação à data de vencimento, existe a carência, data em que um investimento pode ser sacado, por vontade do investidor, mantendo as taxas antes acordadas. Apesar dos títulos apresentarem data de vencimento e carência, alguns podem ser negociados e vendidos no mercado secundário, o qual tem como função a transação dos títulos entre os investidores, e não entre os emissores (ASSAF NETO, 2012), gerando liquidez para o mercado. Porém, a venda antes do prazo pode causar rentabilidade menor do que a esperada, com uma taxa menor do que a acordada para a data final do investimento.

Além disso, os investimentos em renda fixa possuem valor mínimo inicial para investir, como R\$ 1.000,00 e 5.000,00, por exemplo. Assim, como o prazo de vencimento, o valor inicial também está atrelado à taxa de rentabilidade do título, no geral, investimentos com aporte inicial com um valor maior, possui uma maior taxa de juros definida.

Em relação às garantias para certos tipos de investimentos, conforme o Banco Central, em casos de declaração de falência de uma instituição financeira, é acionado o Fundo Garantidor de Crédito (FGC), que tem como objetivo ressarcir o investidor e não poluir o mercado financeiro com um caso isolado. É uma entidade privada, sem fins lucrativos, que administra um mecanismo de proteção aos correntistas, poupadore e investidores, permitindo recuperar os depósitos ou créditos mantidos em instituição financeira, até determinado valor, em caso de intervenção, de liquidação ou de falência. Porém, é importante salientar que não são todos os investimentos que são garantidos pelo FGC.

O total de créditos garantidos pelo FGC por pessoa, contra a mesma instituição financeira, será garantido até o valor de R\$ 250.000,00 (duzentos e cinquenta mil reais) e o total dos créditos de cada investidor contra o conjunto de todas as instituições associadas, será garantido até o valor de R\$ 1.000.000,00 (um milhão de reais) a cada período de 4 (quatro) anos consecutivos.

Em relação à tributação, tem-se o Imposto sobre Operações Financeiras (IOF) e o Imposto de Renda (IR). O IOF em renda fixa é regressivo, ou seja, diminui à medida que aumenta o prazo de aplicação, sendo zero quando o investimento atinge 30 dias. Já o Imposto de Renda depende em qual tipo de investimento está sendo feito, podendo ser isento para pessoas físicas, porcentagem fixa ou implicar com base na tabela regressiva, no qual quanto mais tempo ficar nesse investimento, menor será o imposto a ser cobrado. A tabela regressiva, divulgada pela AMBIMA (2019), se porta da seguinte forma:

- 22,5% para aplicações com prazo de 180 dias;
- 20,0% para aplicações com prazo de 181 até 360 dias;
- 17,5% para aplicações com prazo de 361 até 720 dias;
- 15,0% para aplicações com prazo superior a 721 dias.

No qual, tal alíquota, será aplicada no ganho do investimento e o imposto de renda será retido na fonte no momento do resgate ou vencimento da aplicação.

2.1.2. Tipos de produtos

Os tipos de produtos mais comuns no mercado financeiro são:

a) Certificado de Depósito Bancário

É um título emitido por bancos para captar dinheiro e financiar suas atividades, no qual em troca de tal empréstimo o banco irá devolver ao investidor, após o tempo estipulado, a quantia aplicada e mais o juro que foi acordado no momento do investimento (XP INVESTIMENTOS, 2018).

Existem três tipos de CDB: pré-fixado, pós-fixado e o misto, sendo que a diferença entre eles é como a rentabilidade é calculada, segundo Fk Partners (2018):

- CDB Pré-fixados: a taxa de juros já é definida no momento da aplicação, assim é possível saber no início do investimento qual será o rendimento final do título. Por exemplo: CDB com taxa de 14%;

- CDB Pós-fixado: a rentabilidade é determinada por percentual sobre um índice, como o Certificado de Depósito Interbancário (CDI). Como este índice é variado, a rentabilidade só será definida no dia do vencimento do título. Por exemplo: CDB com taxa de 110% do CDI;

- CDB Misto: é a junção do CDB pré-fixado com o pós-fixado: parte da rentabilidade é estabelecida no momento da aplicação e a outra parte fica relacionada a um índice econômico, como o IPCA. Por exemplo: CDB com taxa de IPCA + 5,5%.

Os investimentos em CDB apresentam a segurança do Fundo Garantidor de Crédito (FGC) e o IOF e IR são calculados através da tabela regressiva na qual o IR incide somente sobre o lucro e no momento de resgate da operação.

b) Letra de Câmbio

Conforme a XP Investimentos (2019), letra de câmbio é um instrumento de captação das financeiras, com o objetivo de financiar suas atividades. Segue as mesmas propostas dos CDB, em que se empresta dinheiro para instituições bancárias, mas nesse caso o empréstimo é feito às financeiras em troca de uma rentabilidade preestabelecida. Desta forma, o investidor empresta dinheiro à instituição financeira, ela utiliza esse valor para financiar empréstimos dos seus clientes e, no fim, a instituição financeira devolve o dinheiro com os juros embutidos do período que foi investido. Sua rentabilidade pode ser pré-fixada, pós-fixada ou

mista, conta com garantia do Fundo Garantidor de Créditos (FGC) e no quesito tributação, é cobrado IOF em até 30 dias e segue a alíquota da tabela regressiva.

Nesse tipo de investimento as taxas de juros pagas podem ser maiores que em relação aos CDB, porém podem fornecer um risco maior de crédito. Geralmente os prazos de vencimento desse investimento são mais longos.

c) Letra de Crédito Imobiliário e Letra de Crédito de Agronegócio

Letra de Crédito Imobiliário (LCI) e a Letra de Crédito de Agronegócio (LCA) são papéis emitidos pelos bancos, com o intuito de arrecadar recursos para a realização de seus empréstimos para o ramo imobiliário e agrícola. Conforme os investimentos já apresentados anteriormente, podem ser pré-fixados ou pós-fixados.

Como esse tipo de investimento é incentivado pelo governo, não sofrem incidência de imposto de renda, tendo garantia do FGC. Porém, pode ser desvantajoso, em relação aos prazos, que costumam ser longos para a realização de saque final (COMO INVESTIR, 2019).

d) Debêntures

Debênture, segundo definição da Bolsa Brasileira intitulada por B3 (Brasil, Bolsa, Balcão) (2019) é um título de dívida que gera direito de crédito ao investidor e são emitidas por sociedades anônimas de capital aberto ou fechado, sendo que o recurso captado é utilizado para financiar projetos ou reestruturar dívidas da empresa. Assim, como os outros títulos de renda fixa, o investidor terá o direito de receber remuneração periodicamente ou no vencimento do título, em que também receberá de volta o valor investido (principal).

As debêntures não contam com o FGC, sendo que o risco está diretamente relacionado ao não pagamento por parte da empresa emissora. O IOF e IR são calculados através da tabela regressiva, onde o IR incide somente sobre o lucro e no momento de resgate da operação.

Uma diversificação em relação as debêntures são as incentivadas ou de infraestrutura, que buscam, conforme a XP Investimentos (2019) afirma, financiar atividades ligadas à infraestrutura, como construção de estradas, linhas de transmissão de energia e saneamento básico, por exemplo. A diferença desse tipo de debênture é a isenção de IOF e IR para pessoas físicas.

e) Títulos Públicos

São títulos utilizados pelo governo para adquirir dinheiro e assim realizar atividades nas áreas de infraestrutura, educação e saúde, além de ser um mecanismo monetário, o qual é possível controlar o dinheiro que circula pela economia do país. São negociados pela plataforma do Tesouro Direto, programa desenvolvido pelo Governo Federal, fusão entre a B3 e o Tesouro Nacional, com o objetivo de funcionar como facilitador para quem quer investir em títulos públicos. Porém, para ter acesso à plataforma do Tesouro Direto é preciso ter uma conta em uma Corretora ou Distribuidora de Valores Mobiliários. Assim como a poupança, é um investimento popular entre os brasileiros e estes são considerados alguns dos títulos mais seguros do mercado devido ao seu emissor, o Governo Federal.

Dentro do tesouro direto existe uma vasta opção de títulos, o que possibilita se adequar ao perfil do investidor e aos seus objetivos, os quais podem ser de curto ou longo prazo, segundo Fk Partners (2018):

- Tesouro Pré-fixado: é um título que tem sua rentabilidade estabelecida, ou seja, fixada já no início da aplicação;

- Tesouro Pré-fixado com juros semestrais: existem as NTN-F (Notas do Tesouro Nacional - Série F), assim como no LTN (Letra do Tesouro Nacional), a rentabilidade é fixada no momento da compra do título. Suas diferenças nos seus pagamentos estão em relação a períodos, onde a cada seis meses o investidor recebe os juros referentes a esse investimento, chamado de cupom. Em relação ao pagamento do dinheiro investido no início, é recebido somente ao fim da aplicação;

- Títulos Pós-fixados ou Letra Financeira do Tesouro (LFT): contrário ao pré-fixado, a rentabilidade é estipulada por meio de um percentual determinado sobre uma taxa, como o IPCA (Índice Nacional de Preços ao Consumidor Amplo) ou o CDI;

- Tesouro IPCA + ou NTN-B Principal (Nota do Tesouro Nacional - Série B Principal): é um investimento que segue à inflação mais a taxa prefixada, como o nome mesmo indica. Essa modalidade traz proteção ao investidor no quesito variação inflacionária, tendo sempre uma rentabilidade real, ou seja, uma rentabilidade acima da inflação do período;

- Tesouro IPCA com juros semestrais ou Nota do Tesouro Nacional – Série B (NTN-B): características próximas ao Tesouro IPCA +, com inflação e soma de uma taxa prefixada, porém os juros rendidos serão pagos a cada seis meses, conhecido como cupom;

- Tesouro Selic: antigamente conhecida como Letra do Tesouro Nacional (LTN) é um título que caminha junto a taxa básica de Juros. É o título mais conservador possível, rendendo sempre 100% da Selic.

Neste contexto, os títulos públicos citados são determinados por data de vencimento, mas apesar dessa data, é possível resgatar antes do prazo, podendo ter um ágio ou deságio do valor investido. Com exceção do Tesouro Selic, título que acompanha sempre 100% da Taxa Selic, não é possível valorizar ou desvalorizar em relação a novas opções do mercado, sendo indicado para ser a reserva de emergência do investidor. Todos os títulos públicos sofrem incidência do IR seguindo a tabela regressiva e IOF nos 30 primeiros dias de aplicação.

f) Ações

Conforme a B3 (2019), a ação é a menor parcela do capital social das companhias abertas ou sociedades anônimas, empresas que são permitidas pela CVM a negociar no mercado de valores mobiliários, ou seja, é um título patrimonial e dá aos seus detentores, que são os acionistas, direitos e deveres de sócio, limitado pela quantidade de ações que possui.

Além disso, ações se encaixam na categoria de renda variável, como o nome já diz: sua renda pode variar e a sua negociação é feita através da Bolsa de Valores, que no Brasil é realizado pela B3, com o preço atualizado em tempo real, com horário pré-determinado e em dias úteis.

Os benefícios que os investidores possuem ao comprar uma ação são: recebimento de dividendos e juros sobre capital próprio, bonificação e direito à subscrição, além disso, o ganho em uma ação pode estar ligado também a valorização do preço da mesma. Dividendos são a distribuição do lucro da empresa no período, após o pagamento do imposto de renda, aos acionistas (ASSAF NETO, 2012). Os juros sobre capital próprio são pagos com base nas reservas patrimoniais de lucro, resultado de exercícios anteriores e que ficaram retidos na empresa. A bonificação é um direito do acionista em receber ações, proporcionais aos títulos possuídos, em decorrência do aumento de capital de uma empresa. Por fim, o direito de subscrição é aquele em quem o acionista tem o direito de ser consultado, em caso de aumento de capital da empresa.

Em relação à alíquota de tributação, fica-se condicionado:

- 15% do total recebido com Juros sobre Capital Próprio (JCP);
- 20% sobre o lucro, em operações *day trade* ou seja, compra e venda no mesmo dia;
- 15% sobre o lucro, para vendas mensais acima de R\$20 mil.

Além do IR, os custos envolvendo uma compra ou venda de ações é a taxa de corretagem, cobrada pelas corretoras e taxa de emolumentos, cobrado pela B3.

O investidor que detêm cotas de empresas listadas na bolsa possui algumas vantagens, como um potencial de maior rentabilidade ao longo prazo, recebimento periódico de dividendos, liberdade de vender ou comprar a qualquer momento, assim como emprestar essas ações e ganhar um rendimento sobre ela (B3, 2019).

2.2. Fundos de Investimentos

Segundo o Banco Central (2019), fundo de investimentos é uma comunhão de recursos constituída sob a forma de condomínio, que tem como objetivo a aplicação em ativos financeiros no mercado financeiro de capitais. Desta forma, é uma aplicação financeira, em que há a combinação de recursos captados de pessoas jurídicas ou físicas, com o propósito de adquirir ganhos financeiros com aplicação em títulos e valores mobiliários. A remuneração varia de acordo com o rendimento dos ativos financeiros que fazem parte do fundo sendo que, não é garantido o recebimento de valor superior ao aplicado, considerando a oscilação que pode existir dos ativos.

Esses ativos financeiros podem variar conforme a classificação e estratégia do fundo. Sendo assim, nas carteiras dos fundos de investimentos pode se encontrar títulos de renda fixa como o CDB, debêntures, títulos públicos e renda variável, como ações, por exemplo.

Para entender o mecanismo dos fundos, existe uma comparação com condomínios, conforme explica Vérios (2019): em fundos de investimentos, cada condômino é dono de uma cota (um apartamento) e paga a alguém (síndico ou administrador) para administrar e coordenar as diversas tarefas do condomínio (limpeza, porteiro etc.). Assim como um condomínio, o fundo de possui um regulamento, com todas as regras de funcionamento e que se aplicam para todos os cotistas. Ao comprar cotas de um determinado fundo, o cotista aceita as regras de funcionamento (aplicação, resgate, custos, entre outros), e passa a pagar uma taxa de administração para que um administrador coordene o funcionamento do fundo.

Os recursos dos investidores de fundos de investimentos, conhecidos como cotistas, são utilizados para comprar títulos que são de todos os investidores na proporção do valor investido. Uma cota é uma fração de um fundo, a soma de todas as cotas resulta no valor do patrimônio do fundo. Sendo que o rendimento da aplicação se dá pela alteração diária do valor da cota (PORTAL DO INVESTIDOR, 2019).

Ao investir em um fundo, o número de cotas do investidor se altera somente quando é aplicado mais recurso, resgatado algum valor ou quando incide a tributação de IR chamada de “come-cota” e, nos outros períodos, a quantidade de cotas fica estática, variando somente o valor unitário dessas cotas.

A administração dos fundos é feita por gestores profissionais, credenciados pela CVM, especializados e que avaliam as melhores opções de investimento, fazendo toda a aplicação e estudo dos investimentos. Também são responsáveis por: gerar rentabilidade e controlar o risco da carteira; acompanhar diariamente os recursos do fundo; avaliar as opções existentes, cenários, acontecimentos políticos e econômicos; e possuem poderes para tomar decisões de investimento com o dinheiro dos cotistas, respeitando o regulamento do fundo (VÉRIOS 2019).

Para pagar a administração é cobrada algumas taxas como a Taxa de Administração, cobrada pela instituição financeira com o objetivo de remunerar os serviços prestados de administração do fundo, da gestão da carteira e distribuição (ASSAF NETO, 2012). Essa taxa se dá por um porcentual anual fixo, aplicado diariamente sobre o patrimônio do investidor. Também existe a Taxa de Performance, cobrada por alguns fundos para remunerar o gestor, devido um bom trabalho. É expressado percentualmente sobre os ganhos que o fundo tiver acima do acordado, por exemplo, 20% acima do CDI ou do Ibovespa. Caso o fundo tenha resultados acima do CDI ou Ibovespa, esse rendimento acima será taxado pela porcentagem combinada. Segundo Fortuna (2010), também podem existir as taxas de ingresso, cobrada na aquisição das cotas ou quando é solicitado o resgate das mesmas e a taxa de saída antecipada, que é quando existe um tempo mínimo de permanência para fazer o resgate, mas o investidor não espera esse tempo.

Os fundos de investimentos são importante alternativa de investimento para pessoas que buscam participar do mercado de capitais, conforme Assaf Neto (2012) afirma, pois traz benefícios como baixo um volume inicial para aplicação, risco minimizado devido as diversas aplicações da carteira do fundo e pelo fato da

administração profissional por parte dos gestores, sem que seja preciso ter um conhecimento profundo sobre investimentos.

Assim como em outras aplicações, nos fundos de investimento, quanto maior o retorno, maior será o risco. Os riscos encontrados nos fundos definidos pelo autor Assaf Neto (2012) são os riscos de crédito, de mercado, liquidez e sistêmico. O risco de crédito envolve a possibilidade de uma emissora dos títulos adquiridos pelo fundo não arcar com suas responsabilidades e não pagar, ou pagar atrasado, refletindo diretamente na rentabilidade geral do fundo.

A parte em relação ao risco de mercado se dá pelas variáveis do mercado, como a taxa de juros da economia, variação cambial, inflação etc., as quais podem interferir na rentabilidade dos títulos e operações da carteira. O risco de liquidez corresponde à dificuldade em vender os títulos, seja por baixa liquidez devido a pouco recurso por parte do mercado ou falta de atratividade. O risco sistêmico se dá pelo comportamento da conjuntura nacional e internacional que interferem nas taxas de juros do mercado, preços, câmbio, por exemplo, interferindo na rentabilidade do fundo como um todo, sendo este risco imprevisível devido às suas variantes.

Outros termos importantes para os Fundos de Investimentos são listados no Quadro 1:

Quadro 1 – Outros termos em Fundos de Investimentos.

Aplicação mínima	Valor mínimo para o primeiro investimento no fundo
Aporte	Quando o cotista investe recursos no fundo. O fundo emite cotas correspondentes ao valor investido.
Come-cotas	Sistema de tributação aplicado a alguns tipos de fundos de investimento.
Cotização	Processo de troca de dinheiro por cotas ou cotas por dinheiro. Na maioria dos fundos, a cotização é diária, atualizando diariamente o rendimento da aplicação
Prazo de Resgate	Prazo estabelecido para que os recursos solicitados sejam depositados em conta. É a soma do prazo de cotização de resgate com o prazo de liquidação de resgate
Movimentação mínima	Valor mínimo exigido para realização de aportes e resgates
Prazo de cotização	É a quantidade de dias que o fundo demora para converter os aportes em cotas (prazo de cotização de aporte) ou para converter cotas em dinheiro (prazo de cotização de resgate).
Prazo de liquidação	É a quantidade de dias que o fundo demora, depois da cotização, para depositar na conta corrente do cotista o valor do resgate efetuado.
Resgate	Quando o cotista retira recursos do fundo. Suas cotas são eliminadas, e o fundo devolve o dinheiro ao cotista.
Investidor qualificado	Pessoas naturais ou jurídicas que possuam investimentos financeiros superior a R\$ 1 milhão de reais e declarem-se qualificadas ou que possua alguma certificação que a CVM aceite considerar como investidor qualificado.
Investidor profissional	Pessoas naturais ou jurídicas que possuam investimentos financeiros superior a R\$ 10 milhões e declarem-se investidores profissionais.

Fonte: Vérios (2019)

Na parte da tributação, na maioria dos fundos é adotada a tabela regressiva de IR sobre o rendimento da aplicação e o IOF é isento em investimentos acima de 30 dias. Além disso, com exceção dos fundos de ações, é recolhido IR antecipado no último dia útil de maio e novembro (“come-cotas”). Em vez de recolher o IR do rendimento da aplicação, o valor devido de IR é transformado para cota, reduzindo as cotas que o investido possui do fundo. A alíquota utilizada para calcular o come cotas é de 20% para fundos de Curto Prazo e 15% para Longo Prazo (FK PARTNERS, 2019).

2.2.1. Tipos de Fundos de Investimentos

Há diversos tipos de fundos, como os que investem em ações, em renda fixa, fundos DI e os fundos multimercados. Os aspectos estruturais de funcionamento são os mesmos, mas cada um tem uma diretriz de investimento própria, chamada política de investimento. Essa política é a que classifica os fundos, conforme regra da CVM e norteia o trabalho que o gestor deve fazer.

a) Renda Fixa

É investido no mínimo 80% do patrimônio líquido em ativos de renda fixa, sendo expostos à variação da taxa de juros ou um índice de preços. A outra parte pode ser aplicada em ações ou títulos com taxa de retorno variável. São compostos por aplicações em título que tem taxa de retorno fixa na maioria das vezes (ABREU, 2019).

Nos fundos de Renda Fixa não há taxa de *performance*, a não ser que seja fundo para investidores qualificados. Para o IR, é seguida a tabela regressiva e IOF até 30 dias, além de incidir o “come-cotas” nos meses de maio e novembro.

b) Renda Fixa de Curto Prazo

Fundos que tem como objetivo ter a menor volatilidade possível dos fundos de investimentos brasileiros. Esse tipo de fundo aplica os recursos exclusivamente em títulos públicos ou privados, prefixados ou indexados à taxa Selic ou outra taxa de juros (ABREU, 2019).

c) Renda Fixa Referenciado

Fundo em que sua política assegure ao menos 95% do patrimônio líquido aplicados em ativos que acompanham, seja diretamente ou indiretamente, o índice de referência determinado, em que tal índice deve estar presente na sua denominação. Além disso, deve possuir 80% dos ativos em ativos do título público ou ativos financeiros considerados como baixo risco de crédito pelo gestor (ABREU, 2019).

d) Renda Fixa Simples

Devem ter no mínimo 95% (noventa e cinco por cento) do patrimônio líquido investido em títulos públicos ou títulos de renda fixa com risco equivalente aos públicos. O sufixo simples deve ter presente em seu nome (ABREU, 2019).

e) Renda Fixa de Dívida Externa

Esses tipos de fundo devem apresentar 80% no mínimo do seu patrimônio investidos em títulos que representam a dívida externa da União. É preciso incluir em seu nome o sufixo Dívida Externa (ABREU, 2019).

f) Ações

É exigido que seja investido no mínimo 67% do patrimônio em ativos de renda variável, sendo que a performance está sujeita a variação dos preços das ações que compõe a carteira, tendo grandes oscilações e é indicado para investimentos com o objetivo de deixar aplicado por um prazo maior (ABREU, 2019). Nesse tipo de fundo o IR é sempre 15% do rendimento, IOF até 30 dias e não possui “come-cotas”.

g) Cambiais

Deve ter como risco principal a flutuação, ou seja, a variação da moeda estrangeira ou do cupom cambial. É preciso manter pelo menos 80% do patrimônio aplicado em ativos relacionados direta ou indiretamente ao fator de risco que dá nome ao fundo (ABREU, 2019).

h) Multimercados

Fundos multimercados são fundos mais amplos e devem possuir política de investimento que envolva vários fatores de risco, sem a necessidade de concentração em determinado fator, ou seja, é possível investir em renda fixa, câmbio, dívida externa e outros. Esse tipo de fundo segue a tabela regressiva para IR, IOF e pode

ser cobrado a taxa de performance, além de ser incidido pelo “come-cotas” (ABREU, 2019).

i) Índice

Os recursos são destinados a aplicação em carteiras que visam refletir a rentabilidade de um índice de referência, por tempo indeterminado. Deve manter no mínimo 95% (noventa e cinco por cento) do seu patrimônio aplicado em ativos de renda variável, na proporção que integram o índice de referência (ABREU, 2019).

j) Direitos Creditórios

Conhecidos por FIDC (Fundo de Investimento em Direito Creditórios), são fundos formado por no mínimo 50% de títulos de crédito originário de operações financeiras realizadas nas instituições financeiras, indústria, arrendamento mercantil, hipotecas, prestação de serviço e outros títulos que são admitidos como direito creditório pela CVM. Neste caso, as operações de crédito são cedidas aos Fundos, que assumem esse direito creditório. Esse tipo de aplicação está disponível somente para investidores qualificados (ABREU, 2019).

k) Fundo de Investimento em Cotas de Fundos de Investimentos

Esse tipo de fundo investe diretamente em outros fundos de investimentos, como o próprio nome diz. Deve manter no mínimo 95% investido em cotas de fundos de investimentos de sua mesma classe, com exceção dos classificados como multimercados, que podem investir de maneira distinta (ABREU, 2019).

Além da regra da CVM, de acordo com a ANBIMA (2019) existem outras classificações que estão relacionadas com características mais específicas, mas para esse trabalho não serão citadas.

Como observa-se pelas classificações dos fundos de investimentos regularizados pela CVM, o mercado é amplo e com isso pode levar ao investidor mais opções de onde pode aplicar, escolhendo conforme for o mais adequado ao seu perfil e objetivo dos seus investimentos.

2.3. Mineração de Dados

O termo mineração de dados faz alusão ao processo de extração de minerais de uma mina, em que a base de dados é a mina, explorada através do uso de

algoritmos, que seriam as ferramentas, a fim de obter conhecimentos específicos, como os minerais preciosos, conforme Nunes e Ferrari (2012).

Para Berry e Linoff (2004), o conceito de mineração de dados pode ser descrito como a exploração e análise de dados em busca de padrões e regras significativas. Para detalhar a exploração e a base de dados, Nunes e Ferrarei (2012) explicam que os dados são símbolos não estruturados e sem significados, como os valores de uma tabela; e a informação está presente nas descrições, a qual apresentam significado e trazem utilidade aos dados apresentados e para finalizar, se acrescenta o conhecimento, o qual é permitido se tomar uma decisão em relação a tudo o que foi apresentado (figura 1).

Figura 1 – Métodos e ferramentas da mineração de dados

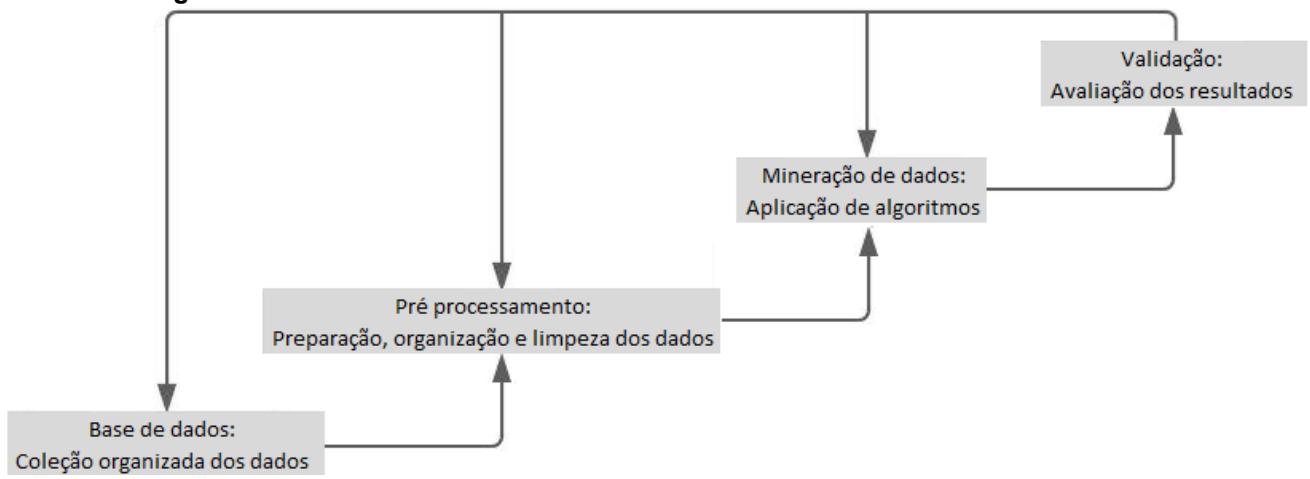


Fonte: Castro e Ferrari (2012)

A mineração de dados é parte integrante do processo de conhecimento, como “Descoberta de Conhecimento em Base de Dados” (*Knowlegde Discovery in Databases*), simbolizado como KDD. Tais conceitos em muitas vezes são compreendidos como sinônimos, mas o KDD se refere a todo o processo de extração de conhecimento a partir de dados e a mineração acontece exclusivamente pela etapa de descoberta do processo de KDD, incluindo a seleção e integração das bases de dados, limpeza da base, seleção e transformação dos dados, mineração e avaliação dos dados (ADRIANNS; ZANTINGE (1996) *apud* CASTRO; FERRARI, 2012).

O processo de KDD pode ser brevemente resumido pela Figura 2.

Figura 2 – Processo de descoberta de conhecimento em base de dados.



Fonte: Castro e Ferrari (2012). Adaptado pela autora

Para a mineração de dados existem diversas tarefas, no qual por meio dessas tarefas é possível extrair diferentes tipos de conhecimentos. Conforme Camilo e Silva (2009) as tarefas mais comuns são:

a) Descrição

Faz parte da etapa inicial do processo de mineração, onde com o uso de ferramentas é possível medir, explorar e descrever características que são reveladas pelos dados. A Descrição é muito utilizada em conjunto com técnicas de análise exploratória de dados, a fim de comprovar a influência de variáveis específicas no resultado que foi gerado.

b) Classificação

A tarefa é considerada uma predição, onde é possível construir e usar um modelo para identificar a qual classe pertence certo registro. A classificação é utilizada para prever dados discretos ou categóricos.

c) Estimação ou Regressão

Tarefa também considerada como predição, porém é usada para estimar o valor de um ou mais atributos. É usada para prever valores contínuos.

d) Agrupamento

Tarefa no qual se dá o processo de separar um conjunto de objetos, em grupos de objetos similares. Diferente da classificação, a classe de cada objeto não é

inicialmente conhecida. Considerando que as classes dos dados não são conhecidas inicialmente, tal processo é conhecido como aprendizagem ou treinamento não supervisionado.

e) Associação

Tarefa que tem como objetivo encontrar relações entre as variáveis e não entre os objetos, ou seja, corresponde a identificação de regras de associação que apresentam valores que ocorrem simultaneamente na base de dados. Como as classes dos dados, assim como na tarefa de agrupamento, não são conhecidas inicialmente, é considerado também como aprendizado não supervisionado.

f) Detecção de anomalias

Como na maioria das ferramentas de mineração, dados que não seguem um comportamento ou não possuem características comuns são descartados, esse tipo de tarefa considera tais dados, pois são importantes para o conhecimento final que se deseja obter.

Para as tarefas de classificação e estimação ocorre o aprendizado supervisionado, em que o modelo analisa o conjunto de registros fornecidos, no qual cada registro já possui a identificação à classe que pertence.

2.3.1. Séries Temporais

Os dados que foram apresentados nesse trabalho são oriundos do valor das cotas dos fundos de investimentos e se comportam como séries temporais e, conforme Morettin (2006), um conjunto de observações realizadas e ordenadas em relação ao tempo, se dá pelo nome de série temporal.

A notação utilizada é:

$$T = \{t_1, t_2, t, \dots, t_m\} \text{ de } m \text{ valores ordenados no tempo.} \quad (1)$$

As séries temporais podem ser contínuas ou discretas, diferenciando quando os valores são registrados continuamente ou os valores são medidos em intervalos específicos de tempo, respectivamente. Também pode-se classificar como

estacionária, que são séries que possuem média e variância constante ao longo do tempo e não-estacionária, em que não é possível definir um padrão (CHINO, 2014).

2.3.2. Regressão linear

Para trabalhar com a tarefa de Regressão, é possível utilizar o método de Regressão Linear (RL), a qual irá detectar o coeficiente da reta que minimiza a distância dos dados, mostrando a tendência de tais dados. A Regressão linear pode ser definida como simples ou múltipla, dependendo da quantidade de variáveis explicativas. E, no caso deste trabalho, será utilizada a regressão simples. Segundo Gareth *et al.* (2017), o método apresenta a abordagem em que é possível prever aproximadamente uma resposta quantitativa, através da relação linear que existe entre dois pontos (X, Y). Matematicamente pode ser escrito por:

$$Y \approx \beta_0 + \beta_1 x + \epsilon \quad (2)$$

Em equação 2 pode-se descrever como regredindo Y em função de X . Onde β_0 e β_1 são duas constantes desconhecidas que representam a interseção (no eixo Y) e inclinação da reta (que representa a tendência), conhecidas como coeficiente da reta. O valor ξ é o erro da predição. E após utilizar dados de treinamento, é possível estimar os valores para β_0 e β_1 , mostrando qual a reta mais se adequa aos dados e prevendo os pontos futuros.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad (3)$$

Onde \hat{y} indica uma previsão de Y , com base em $X = x$.

2.3.3. Redes neurais artificiais

Outra técnica a qual é possível aplicar a tarefa de Regressão é através do conceito da Rede Neural Artificial, inspirada nas redes neurais biológicas existentes no cérebro. O corpo humano consegue realizar atividades de alta complexidade, de maneira simultânea, rápida e automática e isso só é capaz graças à estrutura biológica do ser humano, que é centralizada principalmente no cérebro. Essas estruturas de alta capacidade serviram como exemplo para o desenvolvimento das Redes Neurais Artificiais (RNA), que apresentam como objetivo simular a capacidade de aprender do cérebro.

Para entender as RNAs é preciso compreender como o sistema nervoso funciona, por isso, Faceli *et al.* (2011) explicam e descrevem brevemente os principais pontos biológicos desse sistema. O cérebro faz parte do sistema nervoso, um conjunto complexo de células que dão o funcionamento aos seres vivo, presente em todos os vertebrados e na maioria dos invertebrados. A principal unidade do sistema nervoso é o neurônio, célula nervosa que permite dar resposta aos estímulos, transmitindo impulsos nervosos a outros neurônios, células musculares e das glândulas.

Os componentes principais do neurônio são os dendritos, corpo celular e axônio, são explicados, por Faceli *et al.* (2011):

Dendritos: Prolongamentos dos neurônios responsável e especializado na recepção de estímulos nervosos descendentes de outros neurônicos ou do ambiente.

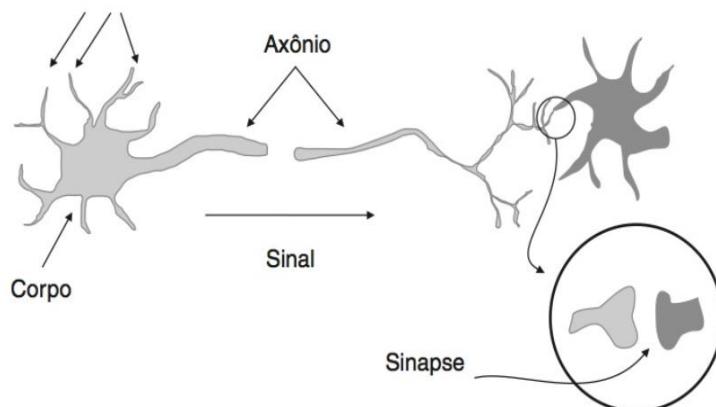
Corpo celular: Recebe os estímulos transmitidos dos dendritos, combina e processa essas informações que foram recebidas. Dependendo da intensidade e frequência do estímulo, cria um estímulo novo, enviando ao axônio.

Axônio: Prolongamento dos neurônios, responsável em conduzir os impulsos elétricos que foram produzidos no corpo celular até outros neurônios.

Sinapse: Unidades que permitem as interações entre os neurônios (HAYKIN (1999) *apud* FACELI *et al.*, 2011).

O sinal do neurônio flui da esquerda para a direita, portanto do dendrito para o corpo celular e posteriormente para o axônio, como representado na Figura 3:

Figura 3 – Representação biológica do neurônio.



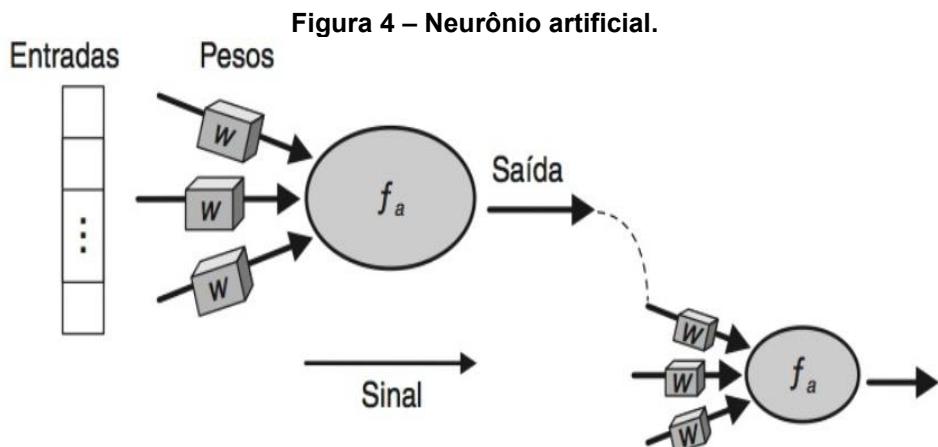
Fonte: Faceli *et al.* (2011)

Para entender a complexidade e capacidade do cérebro humano, Faceli *et al.* (2011) detalha que o cérebro humano possui entre 10 a 500 bilhões de neurônios,

organizados em aproximadamente 1000 módulos principais, cada um com 500 rede neural, acrescentando também que cada neurônio pode estar ligado a centenas ou milhares de outros neurônios. E todas essas redes neurais se organizam de forma paralela e constante, o que gera um processamento muito rápido de informações.

Seguindo o modelo biológico, as RNAs são sistemas computacionais com neurônios artificiais de processamento simples e altamente conectadas. Esses neurônios artificiais açãoam funções matemáticas e estão organizados em uma ou mais camadas, interligados por um alto número de conexões. Nesse caso, as conexões simulam as sinapses biológicas e possuem pesos ligados a elas, que controlam o que os neurônios recebem da rede. Os pesos podem ser valores positivos ou negativos e têm seu valor ajustado no processo de aprendizado, codificando então o conhecimento que foi absorvido pela rede (BRAGA, 2007 *apud* FACELI *et al.*, 2011).

Para entender o funcionamento da RNA, segue na Figura 4, a imagem representativa do neurônio artificial e breve explicação de se seu funcionamento.



Fonte: Faceli *et al.* (2011)

Os terminais de entrada dos neurônios, que simbolizam os dendritos, recebem um valor e esses valores são ponderados e combinados por uma função matemática, representando o papel do corpo celular. A saída da função é a resposta que se obtém para a entrada.

A entrada total recebida por um neurônio (u), pode ser representada pela equação 4:

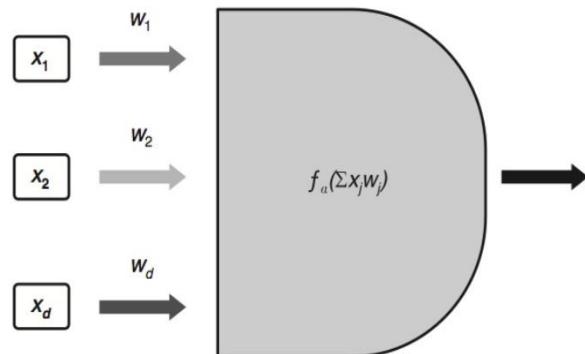
$$u = \sum_{j=1}^d x_j w_j \quad (4)$$

Onde d são os atributos do objeto X , representados em forma de vetor por $x = [x_1, x_2, \dots, x_d]^t$ e um neurônio como d terminais de entrada, com pesos w_1, w_2, \dots, w_d , que podem ser representados vetorialmente por $w = [w_1, w_2, \dots, w_d]$. Gerando a entrada total do neurônio na equação dada acima (equação 4).

Os neurônios podem ter conexões de entradas negativas, positivas ou igual a zero, como $w_j < 0$, $w_j > 0$ e $w_j = 0$ respectivamente. Na conexão que apresenta valor zero para sua entrada há ausência de conexão.

A Figura 5 representa a entrada de um neurônio, onde a função de ativação é definida como a saída do neurônio.

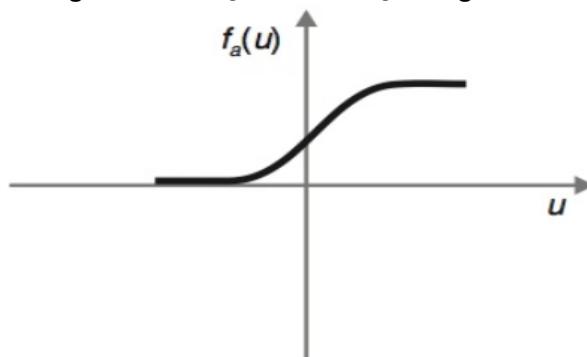
Figura 5 – Representação da entrada de um neurônio artificial



Fonte: Faceli et al. (2011)

A função de ativação que será utilizada nas redes neurais desse trabalho é a função Sigmoidal, onde diferentes inclinações podem ser empregadas, representando uma aproximação contínua e diferenciável.

Figura 6 – Função de ativação Sigmoidal

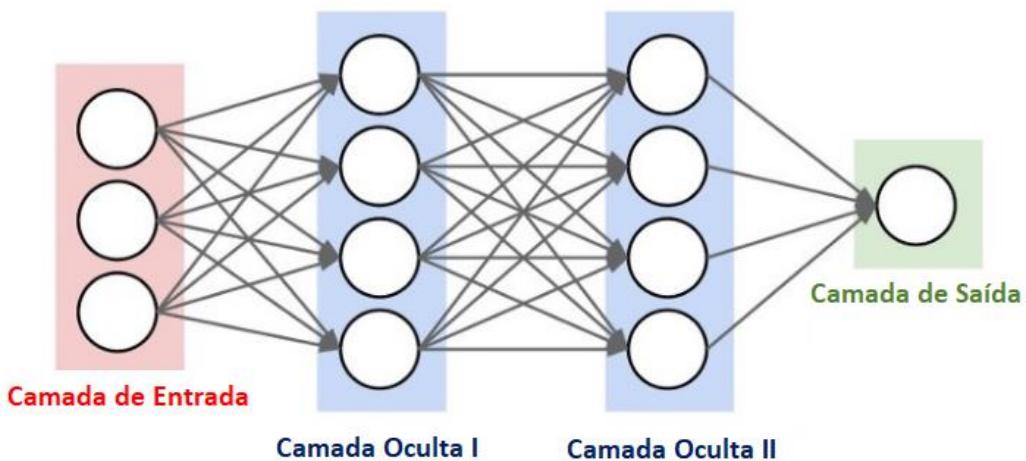


Fonte: Faceli et al. (2011)

Nas redes neurais artificiais, os neurônios podem estar representados em uma ou mais camadas. As Perceptron de múltiplas camadas ou *multi-layer Perceptron* (MLP) são quando duas ou mais camadas são apresentadas. O neurônio pode receber como valor de entrada o valor de saída da camada anterior e/ou passar seu valor de saída para a entrada da próxima camada, até a rede entregar o valor de saída para a última camada. A cada nova camada, a rede fica mais profunda (conhecida como redes neurais profundas ou *deep neural network*) e tem uma capacidade de desenvolver seu aprendizado para situações mais complexas.

Segundo Cybenko (1989) *apud* Faceli *et al.* (2011), redes que possuem camada intermediária, podem implementar qualquer função contínua e é possível fazer a aproximação de qualquer função, pelo uso de duas camadas intermediárias. A Figura 7 exemplifica uma MLP (*Multi Layer Perceptron*) com total de duas camadas, três atributos de entrada e um valor de saída.

Figura 7 – RNA com múltiplas camadas.



Fonte: Gupta (2017). Adaptado pela autora.

Como demonstrado na Figura 7, os resultados serão sempre binários: 1 ou 0, representando a ativação e não ativação, respectivamente. Reforçando o motivo da escolha da sigmoide como função de ativação.

2.3.4. Métrica de avaliação

Para avaliar os resultados encontrados pelos métodos, são utilizadas métricas de avaliação, em que os erros estão baseados nos desvios em relação aos dados reais.

a) Erro Absoluto Médio (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (5)$$

O Erro Absoluto Médio é representado pela equação 5, a qual \hat{y}_i é o valor previsto pelo método, y_i o valor real e n o tamanho da amostra. Segundo Santos (2020), o ponto positivo do método é devido a fácil interpretação do erro pois a unidade de medida avaliada não é alterada.

b) Erro Quadrático Médio (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (6)$$

No Erro Quadrático Médio, o valor residual do MAE ($\hat{y}_i - y_i$) é elevado ao quadrado, conforme representado na equação 6. Para Dos Santos (2020), o MSE trata os dados de forma mais rígida por tratar o erro na forma quadrática, porém se houver pontos de *outliers* nos dados, o erro ficará maior em relação ao Erro Absoluto Médio.

c) Raiz do Erro Quadrático Médio (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (7)$$

A Raiz do Erro Quadrático Médio (equação 7) trabalha com a mesma abordagem do MSE, porém é extraído a raiz de todo o erro. Como os dados ficam na mesma unidade de medida dos dados originais, a interpretação é facilitada e o erro representa o desvio padrão do valor residual (DOS SANTOS, 2020).

Conforme Leandro Castro e Daniel Ferrari (2016), o Erro Quadrático Médio (MSE) e a Raiz do Erro Quadrático Médio (RMSE) tem a capacidade de amplificar grandes discrepâncias entre a saída desejada e o que é previsto.

d) Erro Absoluto Médio Percentual (MAPE)

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (8)$$

O Erro Absoluto Médio Percentual (MAPE), representado pela equação 8, tem como vantagem a capacidade de superar as limitações apresentadas pelo Erro Absoluto Médio, porém apresenta desvantagem quando há dados com valor zero, pois acaba dividindo toda a equação por zero (DOS SANTOS, 2020).

2.4 Trabalhos correlatos

Yan *et al.* (2010), afirmam o Valor Patrimonial Líquido (NAV) previsto de fundos de investimentos selecionado utilizando Rede Neural Artificial, com a utilização do Algoritmo *Backpropagation*. Foram coletados 200 NAV diários, correspondente ao período de 13 de março de 2008 e 31 de dezembro de 2019. Para o treinamento foram utilizados 20 dados para prever o 21º dia e após foi previsto os dados dos dias 180º ao 200º dia. Com o resultado previsto, foi comparado com os dados originais, mostrando que foi possível prever a tendência dos dados e pontos de inflexão da curva.

Mohammad *et al.* (2019), realizaram um estudo com o objetivo de prever os valores de ações utilizando Rede Neural Artificial, a fim de criar carteiras de ações com o melhor retorno e menor risco com os dados previstos, através do Algoritmo Genético (GA) e Modelo de Índice Único (SIM). Foi coletado os dados de 38 de 45 ações que compõem o índice LQ45 da Bolsa de Valores da Indonésia, entre o período de 6 de outubro de 2008 a 21 de setembro de 2018, com um total de 2464 dias. Após a previsão utilizando a RNA, organizou um portfólio de ações considerando os valores previstos através do método SIM e GA, comparando a capacidade de minimizar os riscos e maximizar os lucros dos dois métodos.

No estudo de Wang *et al.* (2012), o objetivo foi prever o Índice de Ações SZII da China, com dados mensais entre o período de 1993 a 2010 totalizando um conjunto de 216 dados, e o Índice *Dow Jones Industrial Average* (DJIA) dos Estados Unidos, entre 1991 a 2010, com um total de 240 dados. São utilizados os métodos de ARIMA (*Autoregressive Integrated Moving Average*), Suavização Exponencial e Rede Neural

Artificial, com a utilização do Algoritmo *Back Propagation* e testado em diferentes proporções no modelo híbrido PHM (*Prognostic and Health Monitoring*) para buscar minimizar o erro se comparado com as técnicas aplicadas de forma individual.

Vieira e Thome (2000) publicaram uma pesquisa que tem como objetivo prever o Índice Bovespa (IBOV) usando Redes Neurais Artificiais, com diferentes critérios de treinamentos e avaliação de desempenho como a acurácia, lucratividade e consistência em relação ao lucro ou prejuízo em negociações, que são métricas importantes para previsões que envolvem finanças.

Foi utilizado o valor de fechamento diário do Ibovespa entre o período de 08 de janeiro de 1996 e 03 de junho de 1998, para as observações e para teste, o período de 4 de junho de 1998 até 08 de janeiro de 1999. Os dados coletados são trabalhados em uma RNA com utilização do algoritmo de *Feedforward* e as previsões foram comparadas entre si, considerando a lucratividade apresentada v.s. acurácia do modelo, consistência de lucro v.s. acurácia e consistência do lucro v.s. lucratividade.

Além das análises para prever ações, fundos e índices, a Rede Neural também pode ser utilizada para prever a inflação de um período. Terra *et al.* (2010) avaliam a eficácia de uma RNA para prever a inflação com o objetivo de auxiliar a gestão de organizações de pequeno porte.

Para a previsão da inflação, foram testados três modelos alterando o número de camada oculta de cada, com os dados mensais entre 1 de janeiro de 1998 e 31 de julho de 2008. Para a variável dependente foi utilizada a inflação, através do Índice IGP-DI (Índice Geral de Preços - Disponibilidade Interna) e as variáveis independentes foram baseadas no trabalho de Pereira e Nakano (1984): Taxa básica de Juros, Taxa de Câmbio, Salários, Resultado Primário, Base Monetária Restrita e Crescimento do Produto Interno Bruto (PIB). Os resultados foram medidos pelo Coeficiente de Determinação (R^2) e erro médio dos modelos, concluindo que é possível utilizar as RNA's para a previsão média da inflação.

No artigo de Xiao e Enke (2016), é apresentado um processo de mineração de dados com o objetivo de prever a direção do retorno diário do Índice S&P 500 Index ETF, considerando 60 características financeiras e econômicas.

Os dados foram coletados considerando a direção do Índice S&P 500 (alta ou queda) de fechamento e 60 fatores financeiros, do período de 1 de junho de 2003 e 31 de maio de 2013, com um total de 2.518 dados diários. É aplicado os métodos de

PCA (*Principal Component Analysis*), FRPCA (*Fast Robust Principal Component Analysis*) e KPCA (*Kernel Principal Component Analysis*) a fim de reduzir a dimensionalidade dos dados. Após isso, foi aplicada a técnica de Redes Neurais, em que os três modelos são comparados, concluindo que a rede neural que foi utilizada a redução pelo PCA gerou precisão maior e que as negociações geraram maior lucro, ajustado ao mesmo risco, se comparando com os dados do benchmark e dos modelos baseadas no FRPCA e KPCA.

Oliveira *et al.* (2008), afirmam a previsão de séries de ações em empresas selecionadas do setor de financeiro, alimentos, industrial e de serviços, com o uso de Redes Neurais e ARIMA (GARCH). Foram utilizadas duas empresas por setor, com os dados diários a partir do dia 03 de janeiro de 2000, totalizando 1.542 dados, sendo 1.300 os dados de treinamento e 242 para a previsão. As técnicas de RNA e ARIMA aplicadas nas ações e os resultados obtidos são comparados por setor, utilizando como medida de desempenho a Raiz do Erro Quadrático Médio (RMSE) e o Coeficiente de Desigualdade de Theil (TIC). Chegando à conclusão de que ambas as técnicas se mostraram eficazes para as previsões pretendidas.

Morales *et al.* (2020) trabalharam com os modelos de Redes Neurais Artificiais, Processos Gaussianos, Tabela de dados, Árvores de decisão, Máquina de vetor de suporte e Regressão Linear, afim de prever o índice de preços e cotações da Bolsa de Valores Mexicana. Foram coletados dados mensais para o período de janeiro de 2019 até dezembro de 2016 da Bolsa de Valores Mexicanas e variáveis determinantes para o Índice, como o Índice *Dow Jones Industrial Average* (DJIA), Índice Nacional de Preços ao Consumidor (NIPC), Índice nacional de preços (CPI), Reservas Internacionais (RI), Rendimentos sobre Certificados do Tesouro (CETES), Taxa de câmbio (USD MXN), Agregado Monetário (M1) e Risco de Default Soberano do México (SDRM). Os desempenhos dos métodos foram medidos através do Erro Percentual Médio Absoluto (MAPE), em que os resultados foram semelhantes, não mostrando uma faixa de erro com variação drástica entre os modelos. Comentado sobre a maior complexidade dos modelos das redes neurais em relação as outras técnicas, mas com pouca diferenciação de erro.

O quadro 2 apresenta um resumo dos trabalhos correlatos a esta pesquisa.

Quadro 2 - Trabalhos correlatos

Autores (ano)	Objetivo	Técnicas de previsão	Métricas
Yan <i>et al.</i> (2010)	Prever o Valor Patrimonial Líquido (NAV) de um fundo de investimento.	Rede Neurais (algoritmo <i>Backpropagation</i>)	MSE.
Solin <i>et al.</i> (2019)	Prever valor de ações e criar uma carteira otimizada de risco v.s. retorno.	Rede neurais; Algoritmo Genético; Modelo de Índice Único (SIM).	MAE; MAPE.
Wang <i>et al.</i> (2012)	Prever Índice de ações SZII e DJIA com um modelo híbrido de técnicas.	Redes neurais (algoritmo <i>Backpropagation</i>); ARIMA; Suavização Exponencial.	ME; RMSE; MAE; MAPE; DA.
Vieira; Antônio (1999)	Comparar diferentes RNA para previsão do Índice Bovespa.	Rede neurais (algoritmo Feedforward).	MSE; Acurácia.
Terra, Leonardo; Passador, Luiz (2012)	Verificar a capacidade de redes neurais auxiliar na previsão da Inflação.	Redes neurais.	R ² ; ME.
Zhong; Enke (2016)	Prever a direção diária de retorno do Índice S&P 500 Index ETF.	Redução de dimensionalidade com PCA, FRPCA e KPCA; Redes neurais.	Teste T pareado; P valeu; Acurácia.
Oliveira <i>et al.</i> (2008)	Prever séries de ações selecionadas do setor financeiro, alimentos, industrial e serviços.	Redes neurais; ARIMA (Garch);	RMSE; TIC.
Morales <i>et al.</i> (2020)	Prever o desempenho da Bolsa de Valores Mexicana (PQI) após a crise de 2008.	Redes Neurais; Processos Gaussianos; Tabela de dados; Árvore de decisão; Máquina de vetores de suporte; Regressão Linear.	MAPE.

Legenda: DA (Acurácia Direcional); MAE (Erro Médio Absoluto); MAPE (Erro Absoluto Médio Percentual); ME (Erro Médio); MSE (Erro Quadrático Médio); RMSE (Raiz do Erro Médio Quadrático); R² (Coeficiente de Determinação); TIC (Coeficiente de Desigualdade de Theil).

Fonte: A autora (2020)

3. METODOLOGIA

Nessa sessão é descrito os conjuntos de dados, formados por três fundos selecionados, a sequência da pesquisa e as ferramentas que foram utilizadas.

3.1 Descrição do conjunto de dados

Foram coletados o valor da cota diária de três fundos de investimentos selecionados, que possuem perfil, classificação e estratégias diferentes.

Os fundos selecionados foram:

a) Western Asset Fundo de Investimento Ações BDR Nível I

O fundo pertence a gestora internacional Western Asset, com sede em Pasadena, Califórnia, e possui mais oitos escritórios entre os continentes do mundo. O fundo tem como objetivo buscar, a longo prazo, valorização compatível com o mercado americano de ações. O objetivo se dá através da aplicação de no mínimo 67% de sua carteira em *Brazilian Depositary Receipts*, classificados como Nível I (BDRs – Nível I), do mercado norte americano de ações. Sendo assim, o investidor fica exposto à variação cambial BRL (Real Brasileiro) x USD (Dólar Americano) e a oscilações e riscos inerentes deste mercado, o que pode exigir dos cotistas uma tolerância para perdas no curto e médio prazo (WESTERN ASSET, 2020).

O fundo adota gestão ativa, baseada em processos de pesquisa e análise fundamentalista de investimentos e de construção da carteira. O processo de seleção de ativos financeiros se baseia na análise de cenários econômicos nacionais e internacionais, o qual avalia as tendências de mercado e condições macroeconômicas e microeconômicas (WESTERN ASSET, 2020).

Na classificação da CVM, o fundo é de Ações, em que deve investir no mínimo 67% da sua carteira em ações e podemos dizer que é um fundo Agressivo, devido à alta volatilidade e oscilação.

b) JGP Strategy Fundo de Investimento em Cotas de Fundos De Investimento Multimercado

JGP (Jakurski, Guedes & Partners) é uma gestora com experiência em investimentos no Brasil e no exterior, fundada em 1998. Possuem equipe sênior e multidisciplinar que faz a gestão de fundos multimercados de com diferentes níveis de

volatilidade, fundos de ações de diferentes estratégias, fundos de crédito que combinam investimentos em títulos de dívida corporativa nos mercados local e internacional e fundos de previdência privada (JGP ASSET, 2020).

Segundo a JGP Asset (2020), o *Strategy* é um fundo multimercado, que investe em ativos financeiros de alta liquidez dos mercados de renda fixa, renda variável e cambial. Tem como objetivo retorno absolutos e superar substancialmente o *benchmark* (CDI) a longo prazo. Possui perfil de multiestratégia, em que são realizados investimentos a longo prazo e de viés mais tático, com uma menor duração, além de operações orientadas pelo cenário macroeconômico e também de perfil fundamentalista. O Fundo visa explorar as melhores oportunidades presentes nos mercados de renda fixa, renda variável, cambial, crédito, commodities e seus derivativos, utilizando, preferencialmente, ativos de alta liquidez.

Na classificação da CVM, é um Fundo de Investimento em Cotas de Fundos de Investimentos, onde investe no mínimo 95% dos seus ativos em cotas do fundo *Master JGP Strategy Master FIM*, de classificação Multimercado, que possui um perfil amplo e sem necessidade de concentração em um único fator. Devido ao perfil e histórico, podemos dizer que é um fundo moderado e possui uma média volatilidade.

c) Daycoval Classic Fundo de Investimento Renda Fixa Crédito Privado

A Daycoval Asset Management é a área especializada em gestão de Fundos de Investimentos do Grupo Daycoval. Foi criada em 2004 e é composta por equipe multidisciplinar que busca avaliar os movimentos do mercado e adaptar a melhor estratégia para os diversos fundos de investimentos que possuem em seu portfólio, como de Renda Fixa, Ações, Multimercado, Previdência e outros (DAYCOVAL ASSET, 2020).

Ainda para a gestora Daycoval (2020), o fundo possui um perfil conservador e tem como objetivo proporcionar rendimentos por meio de investimentos em renda fixa tradicional, investindo em títulos pré e pós fixados, realizando operações que os rendimentos do fundo acompanhem a tendência da variação do CDI. O fundo aplica principalmente em títulos públicos, podendo também investir em títulos privados, onde os emissores foram classificados como baixo risco de crédito, como debêntures, CDB, FIDC e outros fundos.

Na classificação da CVM, é um fundo de Renda Fixa, pois deve investir no mínimo, 80% em ativos de renda fixa. É possível afirmar que é um fundo conservador e de baixa volatilidade. Para efeito de análise, será considerado apenas o valor da cota diária, não considerando os impactos do IR e tributação come-cotas.

A base de dados foi retirada de forma gratuita do site Info Fundos, que apresenta como objetivo fornecer acesso aos principais dados e características de Fundos de Investimentos. Essas informações consistem em dados financeiros qualitativos e quantitativos relacionados ao mercado de fundos (INFO FUNDOS, 2020). No site, foi possível obter a série histórica das cotas dos fundos de investimentos, com os números de cotistas, valor da cota, porcentagem de variação da cota, valor de captação, resgate e patrimônio líquido de dias úteis, padrão de funcionamento do mercado financeiro.

3.2 Sequência da pesquisa

Primeiramente foram selecionados os fundos de investimentos, em que o motivo da escolha se deu ao perfil do fundo e a data de criação dos mesmos, afim de ter um conjunto de dados de vários anos. Com os três fundos selecionados, foi possível obter dados a partir de maio de 2014 para os três fundos, obtendo 6 anos e 5 meses de dados.

Conforme descrito no item 3.1, os dados vieram no formato Planilha do *Microsoft Excel*, o qual foi necessário alterar a formatação da data para o padrão americano (ano – mês – dia) e ajustar a variação da cota, que veio em formato de porcentagem, no qual foi alterada para decimal, mantendo a proporção da porcentagem. Devido a data de criação dos fundos, alguns possuíam mais dados, sendo assim, para padronização dos dados e resultados, os dados anteriores a maio de 2014 foram excluídos, ficando inicialmente dados de 1607 dias úteis para os três fundos, exatamente entre o período de 07/05/2014 e 25/09/2020.

Com a base já editada, os arquivos foram separados, afim de verificar o comportamento do algoritmo com um número de dados diferentes para treinamento:

- Dados de maio de 2014 a setembro de 2020, com o objetivo de ser uma base com mais dados possíveis, chamada de Base completa;

- Dados de setembro de 2019 à 2020, com objetivo de ter 1 ano de dados, chamada de Base um ano;

- Dados de setembro de 2018 à 2020, com o objetivo de ter 2 anos de dados, chamada de Base dois anos.

Após os dados serem ajustados, o arquivo foi alterado para o formato *Attribute Relation File Format* (ARFF), formato padrão do *Weka*. E então, o *software* foi alimentado e os algoritmos testados através da aba *Forecast*. Na execução do algoritmo, foram retirados 20 dados da amostra de teste, ficando com os seguintes números para treinamento e teste:

- Base um ano: total de 253 dados, sendo 233 para treinamento e 20 para teste;
- Base dois anos: total de 504 dados, sendo 484 para treinamento e 20 para teste;
- Base completa: total de 1607 dados, sendo 1587 para treinamento e 20 para teste.

A retirada de 20 dados para teste se dá pelo padrão que foi adotado para esse trabalho, em que um mês possui 20 dias úteis, dividido por quatro semanas, com cinco dias úteis cada.

A utilização de 20 dias de teste para todas as bases de dados é devido a escolha neste trabalho de testar a previsibilidade do método para até um mês, para as diferentes bases de treinamento.

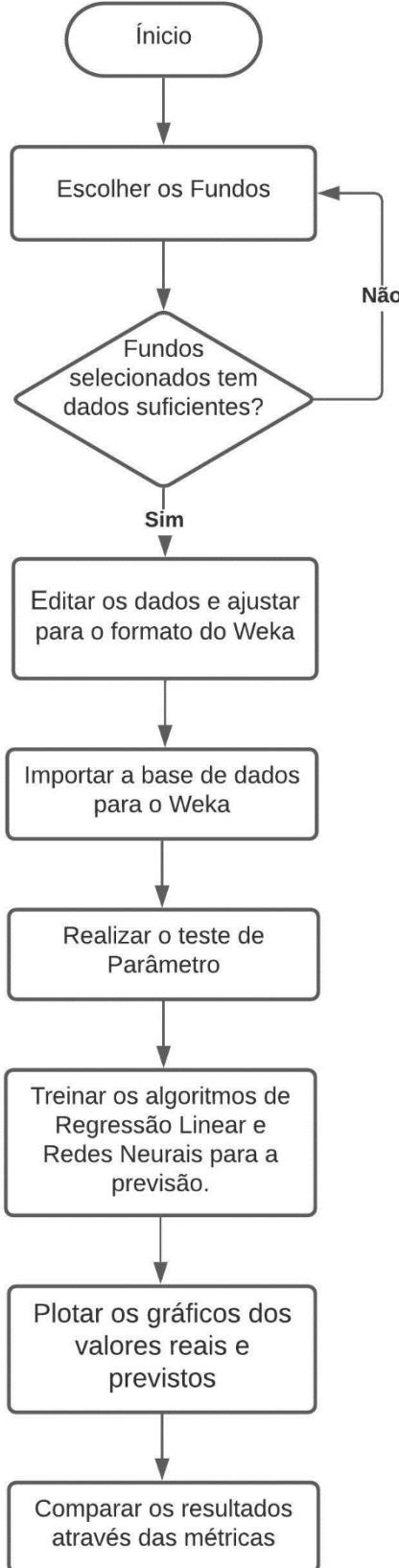
Para iniciar o teste dos algoritmos, foi feito um teste de análise de parâmetro, mais precisamente na Rede Neural, afim de verificar qual taxa de aprendizagem entre 0,1, 0,2 ou 0,3 retorna um menor erro, e por conveniência, o teste foi feito somente para a base completa.

Conforme Park *et. al.* (2020), um dos fatores importantes para a otimização de resultados é utilizar um hiper parâmetro chamado de taxa de aprendizado: se a norma do gradiente for grande, o aprendizado não funcionará bem; se a norma do gradiente é pequena, o aprendizado será muito lento. Após o teste e escolha da melhor taxa de aprendizagem para a Rede Neural, os algoritmos de Regressão Linear e Redes Neurais foram treinados e testados para as três bases, dos três fundos selecionados.

Com os dados treinados, o algoritmo organizou a previsão, plotou os gráficos (comparando os dados reais v.s. previsto) e mostrou os erros de previsão (pelas métricas de MAE, RMSE e MSE), obtendo material para comparar as técnicas e

debater sobre os dados gerados. Na figura 8 é apresentado o fluxograma que resume a sequência da pesquisa realizada:

Figura 8 – Passo a passo da pesquisa



Fonte: A autora (2020)

3.3 Ferramentas utilizadas

Os dados retirados do site Info Fundos foram baixados no formato Planilha do *Microsoft Excel*, formato padrão do programa, e foram ajustados no próprio *Excel* e transformados para formato ARFF. As técnicas de *Machine Learning* e visualização dos dados foram realizadas no *software Weka*.

Waikato Environment for Knowledge Analysis (Weka) é um *software* com coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Ele contém ferramentas para preparação, classificação, regressão, agrupamento, regras de associação de mineração e visualização de dados. É amplamente utilizado em aplicações de ensino e pesquisa, com um grande número de ferramentas integradas para tarefas de aprendizado de máquina padrão e também fornece acesso transparente a caixas de ferramentas conhecidas, como *Scikit-learn*, R e *Deeplearning4j*.

O *software* de código aberto foi emitido sob a *General Public License* (GNU), no qual pode ser baixado no site próprio. Neste, há alguns cursos online gratuitos e os vídeos desses cursos estão disponíveis também no *Youtube*. Para prever os dados foi utilizada a extensão *Forecast* do *software*, a qual é possível baixar a extensão de maneira gratuita dentro do próprio programa.

4. RESULTADOS E DISCUSSÕES

Nessa sessão foram discutidos e avaliados os resultados obtidos para o Teste de Parâmetro da Rede Neural e também as previsões realizadas pelos algoritmos de Regressão Linear e Rede Neural, com os três fundos e respectivas bases de dados.

Os dados foram organizados com o objetivo de prever um mês futuro, por isso foram retirados 20 dias úteis do conjunto de dados para teste. Para a análise dos resultados, as previsões foram plotadas mostrando qual o erro de uma previsão de um dia futuro, cinco, 10, 15 e 20 dias, o que totalizou previsão de até um mês futuro.

4.1 Teste de Parâmetros

Para verificar o melhor parâmetro para os algoritmos, foram feitos diferentes testes com Taxa de Aprendizagem (*Learning Rate*) iniciando com 0,1, depois 0,2 por fim 0,3, aplicado à rede neural. Esse teste foi executado considerando a base completa dos três fundos selecionados.

Com os erros obtidos pela previsão, foi selecionado o melhor parâmetro, para então, ser aplicado no algoritmo de Rede Neural e depois comparado com os resultados obtidos pela Regressão Linear.

a) Fundo Conservador

As tabelas 1, 2 e 3 representam os erros obtidos pelas previsões, utilizando uma taxa de aprendizagem diferente para o fundo Conservador.

Tabela 1 – Teste da Taxa de Aprendizagem 0,1 para o fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0276	0,0342	0,0012
5 dias futuros	16	0,0283	0,0346	0,0012
10 dias futuros	11	0,0307	0,0413	0,0017
15 dias futuros	6	0,0378	0,0438	0,0019
20 dias futuros	1	0,0023	0,0023	0,0000

Fonte: A autora (2020)

Os erros MAE, RMSE e MSE obtidos com a taxa 0,1 são próximos de zeros, observa-se que o erro aumenta conforme os dias previstos aumentam, porém, um melhor resultado é obtido na última previsão (20 dias futuro).

Tabela 2 – Teste da Taxa de Aprendizagem 0,2 para o fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0282	0,0337	0,0011
5 dias futuros	16	0,0272	0,0321	0,001
10 dias futuros	11	0,0295	0,0359	0,0013
15 dias futuros	6	0,0339	0,0397	0,0016
20 dias futuros	1	0,0108	0,0108	0,0001

Fonte: A autora (2020)

Com a taxa 0,2 observa-se o mesmo comportamento da taxa 0,1, de aumento conforme mais dias e depois uma melhor previsão para a última semana, com valores de erros diferentes entre as taxas, porém pouco significante.

Tabela 3 – Teste da Taxa de Aprendizagem 0,3 para o fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0285	0,0335	0,0011
5 dias futuros	16	0,0267	0,0311	0,001
10 dias futuros	11	0,0301	0,035	0,0012
15 dias futuros	6	0,0318	0,0395	0,0016
20 dias futuros	1	0,0269	0,0269	0,0007

Fonte: A autora (2020)

Para a taxa 0,3 o comportamento se mantém o mesmo, com um erro maior que as taxas 0,1 e 0,2, porém ainda de pouca significância.

Devido ao perfil de baixa oscilação do fundo conservador, os erros obtidos pelas três taxas de aprendizado se mostram próximos a zero e não possui diferenças significativas entre si. Sendo assim, por conveniência, foi adotado a taxa de 0,1 para o aprendizado da Rede Neural do fundo Conservador.

b) Fundo Moderado

Seguindo o mesmo para o fundo Moderado, as tabelas 4, 5 e 6 mostram os erros obtidos pelas diferentes taxas de aprendizado.

Tabela 4 – Teste da Taxa de Aprendizagem 0,1 para o fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0784	0,1010	0,0102
5 dias futuros	16	0,0639	0,0762	0,0058
10 dias futuros	11	0,0497	0,0598	0,0036
15 dias futuros	6	0,0585	0,0650	0,0042
20 dias futuros	1	0,0020	0,0020	0

Fonte: A autora (2020)

Os erros obtidos pela taxa 0,1 são baixos e variados conforme os dias previstos aumentam, obtendo um melhor resultado para a última previsão (20 dias futuro).

Tabela 5 – Teste da Taxa de Aprendizagem 0,2 para o fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0963	0,1197	0,0143
5 dias futuros	16	0,0753	0,0879	0,0077
10 dias futuros	11	0,0548	0,0637	0,0041
15 dias futuros	6	0,0475	0,0542	0,0029
20 dias futuros	1	0,0286	0,0286	0,0008

Fonte: A autora (2020)

Para a taxa 0,2 o comportamento dos erros varia da mesma forma que a taxa 0,1, porém possui um erro superior ao primeiro parâmetro.

Tabela 6 – Teste da Taxa de Aprendizagem 0,3 para o fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,1065	0,132	0,0174
5 dias futuros	16	0,0803	0,0947	0,009
10 dias futuros	11	0,0573	0,0685	0,0047
15 dias futuros	6	0,0418	0,0562	0,0032
20 dias futuros	1	0,0454	0,0454	0,0021

Fonte: A autora (2020)

No teste com a taxa 0,3, algumas previsões foram melhores e outras piores em relação a taxa 0,2, porém comparando com o 0,1 os erros foram maiores. Tendo então um pior resultado se comparado com as outras taxas de aprendizado.

Desta forma, a taxa 0,1 se mostrou eficaz em comparação aos outros testes, adotada então como a melhor taxa para o algoritmo.

c) Fundo Agressivo

No teste com o fundo agressivo os erros são maiores em relação ao fundo conservador e moderado, a qual os erros estavam próximos a zero e agora os valores são mais expressivos, no qual se deve a característica do fundo Agressivo possuir uma maior oscilação em relação aos outros. Nas tabelas 7, 8 e 9 é possível visualizar os erros das previsões.

Tabela 7 – Teste da Taxa de Aprendizagem 0,1 para o fundo Agressivo

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	2,0067	2,551	6,5076
5 dias futuros	16	1,9371	2,1744	4,7282
10 dias futuros	11	2,0683	2,4155	5,8347
15 dias futuros	6	1,0638	1,1864	1,4075
20 dias futuros	1	1,5326	1,5326	2,3489

Fonte: A autora (2020)

Os erros para a taxa 0,1 do fundo agressivo são mais expressivos que dos outros fundos, o que já era esperado devido ao perfil do fundo. Nesse caso, a partir da previsão do 15º dia já se observa melhor previsão se comparado aos dias 1, 5 e 10, o que não aconteceu nos fundos conservadores e moderados.

Tabela 8 – Teste da Taxa de Aprendizagem 0,2 para o fundo Agressivo

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	2,5945	3,7025	13,7087
5 dias futuros	16	2,2066	2,5898	6,7071
10 dias futuros	11	3,4204	4,4333	19,6541
15 dias futuros	6	3,2135	5,0324	25,325
20 dias futuros	1	1,1501	1,1501	1,3227

Fonte: A autora (2020)

Para a taxa 0,2 houve melhor previsão para o 20º dia, se comparado ao primeiro teste, porém no geral os erros com essa taxa de aprendizagem foram maiores.

Tabela 9 – Teste da Taxa de Aprendizagem 0,3 para o fundo Agressivo

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	2,6073	3,3751	11,391
5 dias futuros	16	2,566	3,1035	9,6317
10 dias futuros	11	1,9117	2,4734	6,1178
15 dias futuros	6	4,8458	8,1990	67,2243
20 dias futuros	1	1,8697	1,8697	3,4958

Fonte: A autora (2020)

Com a taxa 0,3 não se obteve melhor previsão se comparado com os testes anteriores, principalmente a taxa de aprendizagem de 0,1. Sendo assim, também foi escolhido como hiper parâmetro a taxa de aprendizagem 0,1 para a Rede Neural do fundo Agressivo.

4.2 Regressão Linear

Desta forma, com os parâmetros ajustados da Rede Neural para os três tipos de fundos, iniciou os treinamentos e testes com o algoritmo de Regressão Linear, para as três bases dos fundos selecionados.

a) Fundo Conservador

Inicialmente o algoritmo de regressão foi testado para o fundo conservador, no qual possui um perfil de pouca oscilação (tabela 10).

Tabela 10 – Algoritmo de Regressão Linear aplicado a base de um ano do fundo Conservador

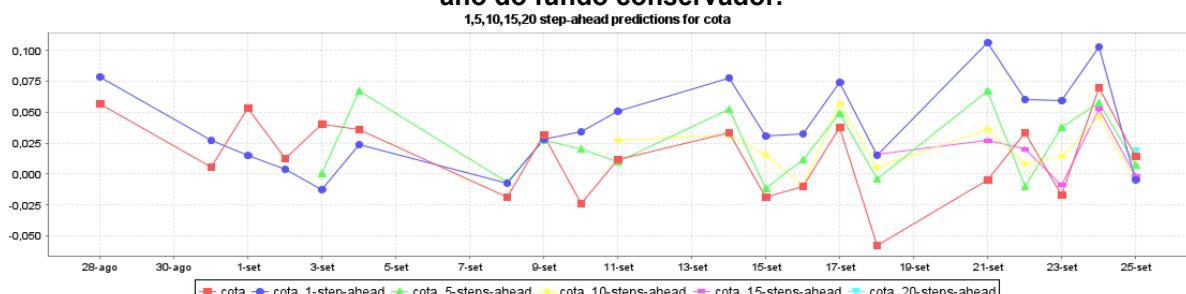
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0390	0,0468	0,0022
5 dias futuros	16	0,0273	0,0344	0,0012
10 dias futuros	11	0,0251	0,0302	0,0009
15 dias futuros	6	0,0268	0,0349	0,0012
20 dias futuros	1	0,0039	0,0039	0

Fonte: A autora (2020)

Com o algoritmo de Regressão observa-se boa previsão para o fundo conservador, com erro próximo do zero em todas as datas prevista. Porém, verificou-se uma assertividade maior em previsões futuras mais longas se comparado com a previsão de um dia futuro: 0,0390 contra um MAE de 0,0039 para a previsão do 20º dia.

No gráfico 9 é possível verificar sobre a assertividade maior para as previsões futuras (linha azul x outras linhas) e também como as previsões acompanharam o comportamento do gráfico real. Apesar dos erros estarem todos próximos de zero, a escala do gráfico pode passar a impressão de que o erro da previsão é maior.

Figura 9 – Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo conservador.



Fonte: A autora (2020)

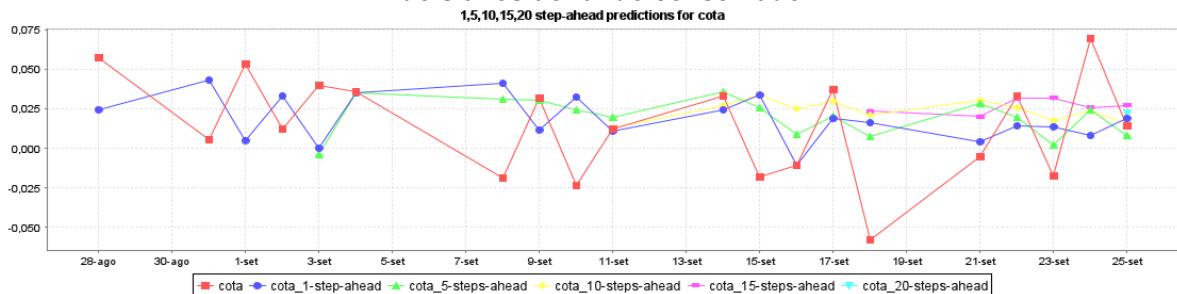
Com a previsão utilizando uma base de dados de dois anos, (tabela 11) observa-se que os erros aumentaram se comparando com a base de um ano, mas mantendo ainda um erro de previsão próximo do zero.

Tabela 11 – Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0298	0,0373	0,0014
5 dias futuros	16	0,0261	0,0329	0,0011
10 dias futuros	11	0,0279	0,0369	0,0014
15 dias futuros	6	0,0358	0,0444	0,0020
20 dias futuros	1	0,0081	0,0081	0,0001

Fonte: A autora (2020)

Figura 10 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo conservador



Fonte: A autora (2020)

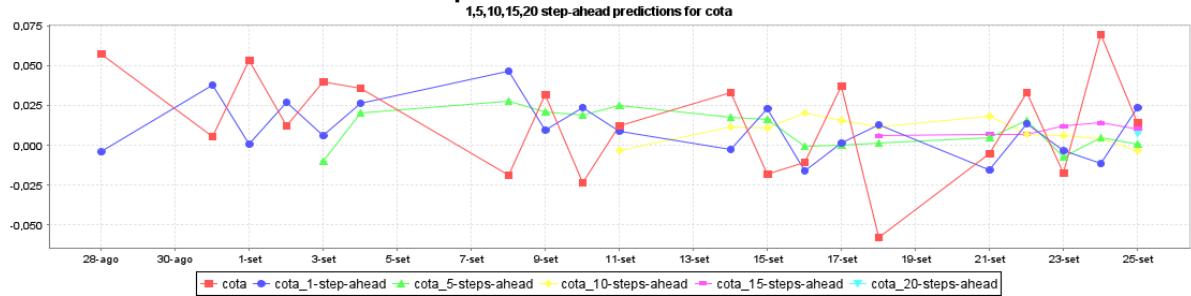
Por meio do exemplo do que aconteceu com a base de dois anos, na base de aproximadamente 7 anos de dados (tabela 12) ocorreu aumento nos erros, porém com erros baixos que se mantém próximos do zero. Desta forma, assim como as outras previsões, o algoritmo apresentou melhor assertividade nas previsões com maior espaço de dias futuros, chegando a 0,0070 para 20 dias futuros contra 0,0333 de 1 dia futuro, para o MAE.

Tabela 12 – Algoritmo de Regressão Linear aplicado a base completa do fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0333	0,0405	0,0016
5 dias futuros	16	0,0281	0,0338	0,0011
10 dias futuros	11	0,0313	0,036	0,0013
15 dias futuros	6	0,032	0,0386	0,0015
20 dias futuros	1	0,007	0,007	0

Fonte: A autora (2020)

Figura 11 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de completa do fundo conservador



Fonte: A autora (2020)

b) Fundo Moderado

Agora serão apresentados os resultados dos erros encontrados ao testar o fundo moderado, fundo com característica uma média oscilação (tabela 13).

Tabela 13 – Algoritmo de Regressão Linear aplicado a base de um ano do fundo Moderado

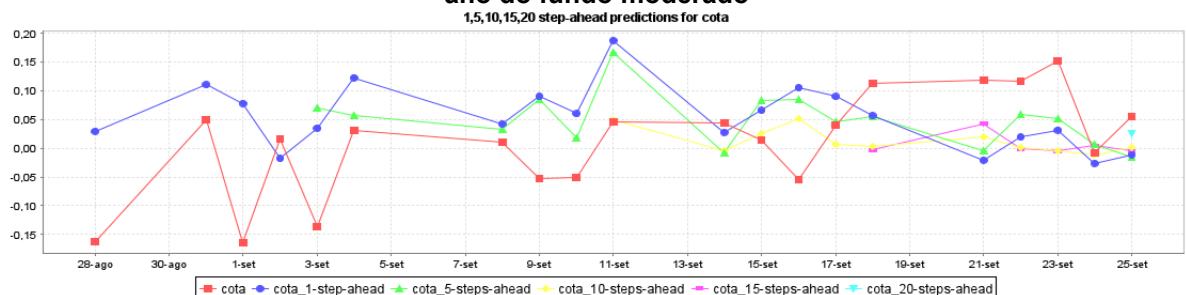
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0997	0,1171	0,0137
5 dias futuros	16	0,0793	0,0954	0,0091
10 dias futuros	11	0,0674	0,0838	0,0070
15 dias futuros	6	0,0896	0,1011	0,0102
20 dias futuros	1	0,0299	0,0299	0,0009

Fonte: A autora (2020)

Ao treinar o algoritmo para o fundo moderado, observa-se que o erro aumentou em relação ao conservador, o que já era esperado devido ao perfil dos fundos. A capacidade de previsão oscila entre os dias futuros previstos, tendo uma melhor resposta na previsão de 20 dias futuros.

Com o gráfico da figura 12 é possível perceber alguns pontos a qual a previsão foi excelente como nos dias 2, 8, 14, 17 e 24 de setembro, por exemplo. Porém em alguns dias, o gráfico mostra um comportamento ao contrário do real, como nos dias 9 e 10 de setembro e 15 e 16 de setembro.

Figura 12 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo moderado



Fonte: A autora (2020)

Para a previsão utilizando a base de dois anos (tabela 14), observa-se uma redução dos erros, diferente do que aconteceu para o fundo conservador. Temos um MAE de 0,075 para a previsão de um dia para frente, contra um erro de 0,0997 no conservador. É possível analisar também essa diferença quando se verifica a previsão de 20 dias futuros em um MAE de 0,0078 e 0,0299 para o fundo conservador.

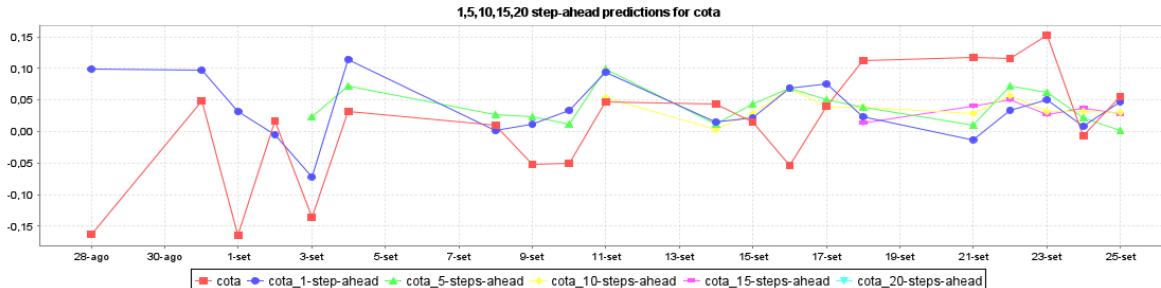
Tabela 14 – Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,075	0,0985	0,0097
5 dias futuros	16	0,0626	0,0742	0,0055
10 dias futuros	11	0,0541	0,068	0,0046
15 dias futuros	6	0,0732	0,0801	0,0064
20 dias futuros	1	0,0078	0,0078	0,0001

Fonte: A autora (2020)

Ao comparar os gráficos de um e dois anos (figuras 12 e 13), observa-se comportamento aparentemente igual, porém o espaçamento entre as retas previstas e o dado real é menor com a segunda base de dados.

Figura 13 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo Moderado



Fonte: A autora (2020)

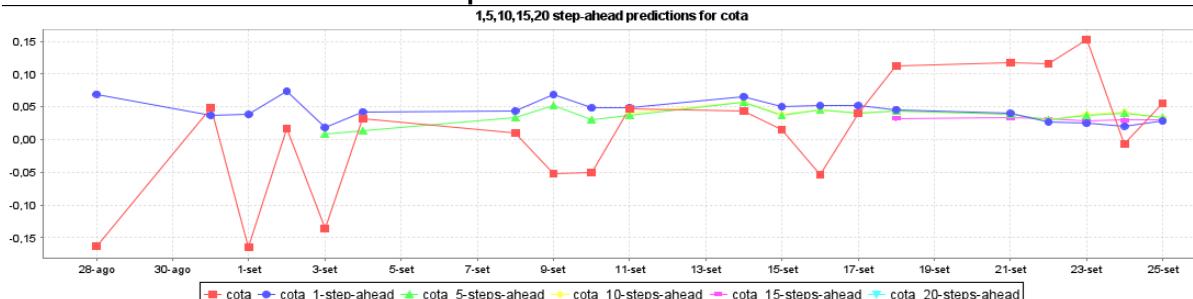
Comparando os dados previstos entre a base completa (tabela 15) e de dois anos (tabela 14), não ocorreram mudanças consideráveis, pois em algumas datas a base de dois anos, o algoritmo teve uma previsão melhor, como em um e 20 dias futuro, por exemplo. Porém as diferenças entre os erros dessas previsões se mostraram baixo, desta forma, não é possível afirmar em qual base de treinamento o algoritmo se mostrou mais eficaz.

Tabela 15 – Algoritmo de Regressão Linear aplicado a base completa do fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0759	0,0994	0,0099
5 dias futuros	16	0,0587	0,0728	0,0053
10 dias futuros	11	0,0512	0,0641	0,0041
15 dias futuros	6	0,0727	0,0799	0,0064
20 dias futuros	1	0,0257	0,0257	0,0007

Fonte: A autora (2020)

Por meio de análise do gráfico da figura 14, os dados reais *versus* previstos (Figura 14), a base completa apresentou comportamento parecido com o de dois anos, assim como foi possível perceber inicialmente ao analisar os erros gerados. Porém, os dados previstos da base completa se aproximaram mais entre si do que no teste anterior, ou seja, na base completa, a previsão sofreu menos o impacto de quantos dias futuros estava prevendo.

Figura 14 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de completa do fundo Moderado

Fonte: A autora (2020)

c) Fundo Agressivo

Devido ao perfil do fundo agressivo, espera-se que os erros de previsão sejam maiores do que dos fundos conservadores e agressivos. A tabela 16 apresenta a Regressão linear para a base de um ano do fundo agressivo.

Tabela 16 – Algoritmo de Regressão Linear aplicado a base de um ano do fundo Agressivo

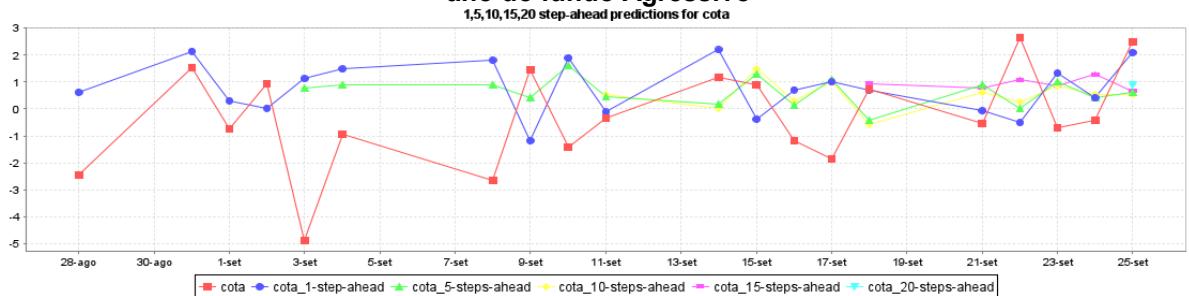
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	1,9343	2,4598	6,0508
5 dias futuros	16	1,9508	2,3434	5,4916
10 dias futuros	11	1,4764	1,6147	2,6072
15 dias futuros	6	1,3666	1,4722	2,1673
20 dias futuros	1	1,5905	1,5905	2,5297

Fonte: A autora (2020)

O erro da Regressão Linear para o fundo agressivo é consideravelmente alto se comparado aos fundos conservador e moderado, onde os erros estavam próximos de zero e agora está entre 1 e 4. O que se observa nessa previsão é que os erros vão diminuindo conforme os dias futuros de previsão aumentam. Passando de um MAE de 1,93 para 1,59, RMSE DE 2,45 para 1,59 e MSE de 6,05 para 2,52.

Com o gráfico comparativo plotado (figura 15) é possível perceber algumas previsões distantes dos dados reais, como no dia 3 de setembro, mas observa-se que entre o dia 11 de setembro e 23 de setembro alguns períodos onde os gráficos tiveram o mesmo comportamento de alta e queda, além como alguns pontos certeiros.

Figura 15 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de um ano do fundo Agressivo



Fonte: A autora (2020)

No teste para dois anos (tabela 17), tirando a previsão para 20 dias futuros, as previsões tiveram um erro menor que da base de um ano. Iniciando o erro com 1,59 e diminuindo para 1,29 no 15º dia, contra 1,93 e 1,36 de MAE em relação à base anterior.

Tabela 17 – Algoritmo de Regressão Linear aplicado a base de dois anos do fundo Agressivo

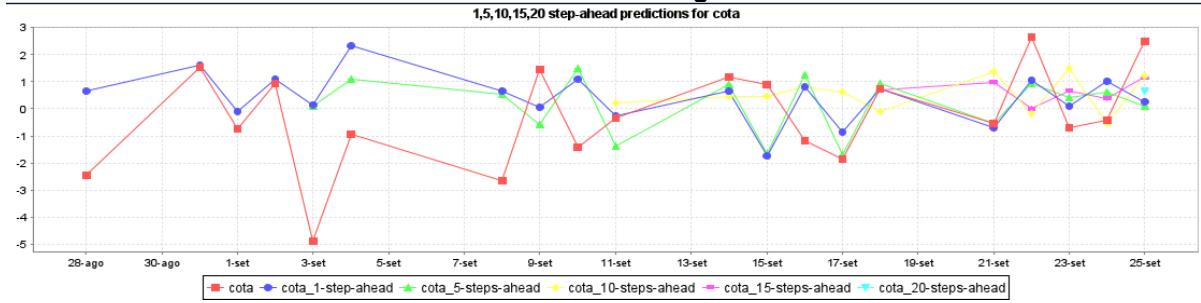
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	1,5942	2,0936	4,3833
5 dias futuros	16	1,7517	2,1843	4,7714
10 dias futuros	11	1,3913	1,6439	2,7023
15 dias futuros	6	1,2907	1,5070	2,2710
20 dias futuros	1	1,8283	1,8283	3,3426

Fonte: A autora (2020)

Pelo gráfico (figura 16) é possível perceber como as previsões e o comportamento entre o dia 11 e 23 de setembro foi ainda eficaz, porém com a previsão do dia 3 de setembro ainda bem distante, podendo entender que houve algum

comportamento no mercado brusco onde o algoritmo ainda não foi capaz de prever, apesar de todo o treinamento.

Figura 16 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base de dois anos do fundo Agressivo



Fonte: A autora (2020)

Para os dados da base completa (tabela 18), os erros não possuíram uma diferença significativa entre a base de dois anos, porém confirmou o mesmo comportamento de ter uma melhor previsão até o 15º dia futuro.

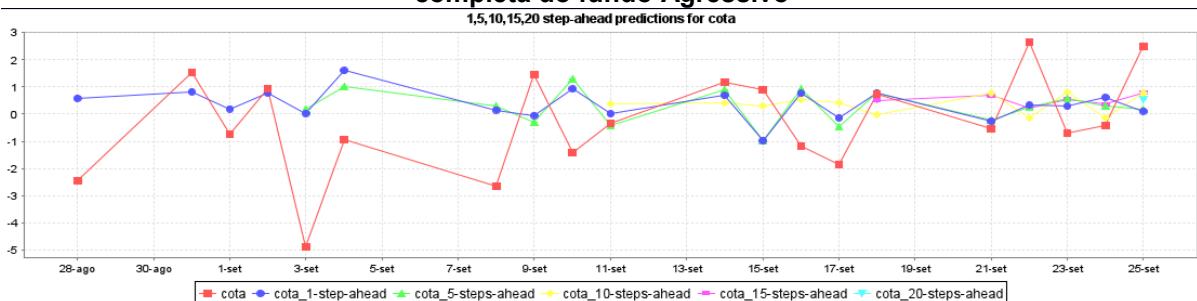
Tabela 18 – Algoritmo de Regressão Linear aplicado a base completa do fundo Agressivo

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	1,6186	2,0073	4,0291
5 dias futuros	16	1,7035	2,1208	4,4979
10 dias futuros	11	1,3161	1,5117	2,2851
15 dias futuros	6	1,2765	1,4500	2,1025
20 dias futuros	1	1,9696	1,9696	3,8794

Fonte: A autora (2020)

No gráfico (Figura 17) entre os dados reais e previstos, observa-se uma posição parecida com o gráfico de 2 anos, porém assim como aconteceu no gráfico do fundo Moderado, a previsão considerando quando dias futuros estava prevendo diminuiu entre sim, entendendo que a base completa sofre menor impacto de quantos dias futuros está prevendo.

Figura 17 - Dados reais e previstos pelo algoritmo de Regressão Linear, aplicado a base completa do fundo Agressivo



Fonte: A autora (2020)

4.3 Rede Neural

Da mesma forma em que os dados foram treinados e testados com o algoritmo de Regressão Linear, serão aplicados para o algoritmo de Rede Neural, utilizando para todos os casos a taxa de aprendizagem 0,1, definido pelo teste de parâmetro realizado.

a) Fundo Conservador

Ao longo serão apresentados como o algoritmo de Rede Neural se comportou para um fundo de baixa oscilação, que são os fundos conservadores. A tabela 19 apresenta a RNA para a base de um ano do fundo Conservador.

Tabela 19 – Algoritmo de Rede Neural aplicado a base de um ano do fundo Conservador

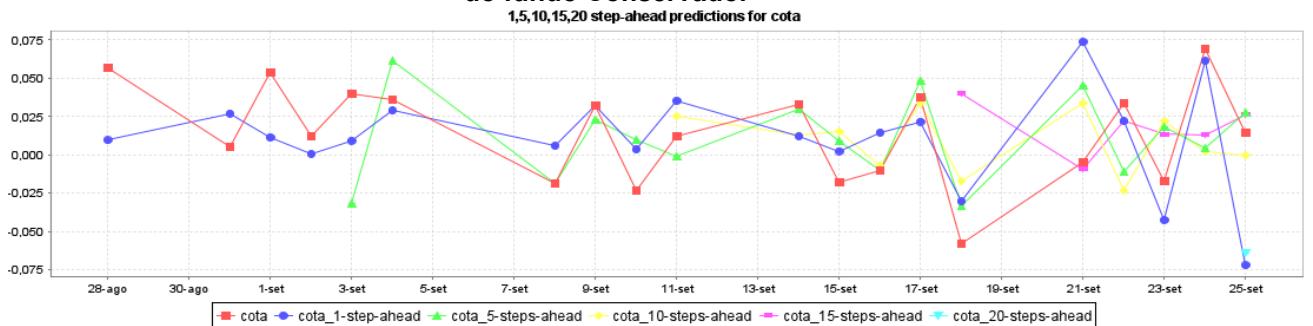
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0278	0,035	0,0012
5 dias futuros	16	0,0267	0,0343	0,0012
10 dias futuros	11	0,0302	0,0362	0,0013
15 dias futuros	6	0,0354	0,0484	0,0023
20 dias futuros	1	0,0784	0,0784	0,0061

Fonte: A autora (2020)

Através da previsão gerada pela Rede Neural, nota-se que conforme os dias futuros de previsão aumenta, os erros também aumentam, mas mantendo erro próximo do zero.

Devido ao erro próximo ao zero e a escala do gráfico (Figura 18), a percepção de erro aparenta ser maior. Porém, é possível perceber como os dados previstos acompanham o comportamento da reta real na maior parte do período previsto.

Figura 18 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Conservador



Fonte: A autora (2020)

Diferente da base de um ano, nessa previsão (tabela 20) os erros não apresentaram perfil exato conforme os dias de previsão aumentaram, variando entre um erro maior e menor. Ainda é possível perceber que a previsão para 20 dias futuros foi a melhor, com um erro muito baixo se comparado as outras previsões.

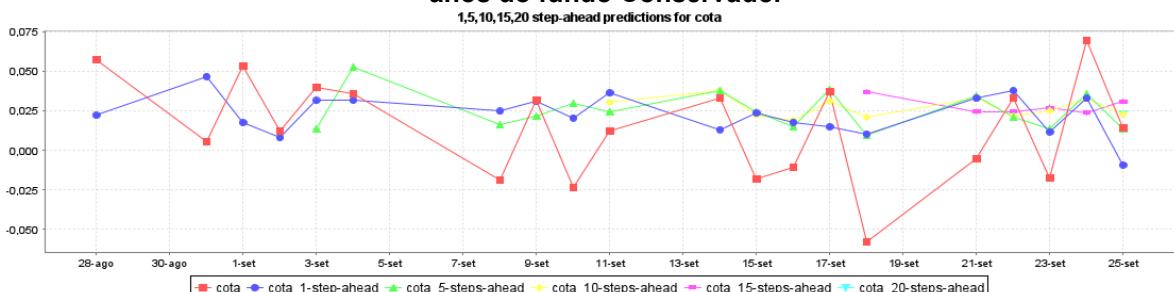
Tabela 20 – Algoritmo de Rede Neural aplicado a base de dois anos do fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0277	0,0325	0,0011
5 dias futuros	16	0,0259	0,0318	0,0010
10 dias futuros	11	0,0288	0,0357	0,0013
15 dias futuros	6	0,0399	0,0487	0,0024
20 dias futuros	1	0,0088	0,0088	0,0001

Fonte: A autora (2020)

É possível perceber pelo gráfico (figura 19) como as previsões não acompanharam um perfil exato, o que gerou um gráfico visivelmente mais bagunçado se comparado ao gerado pelos resultados da base de dados de um ano.

Figura 19 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Conservador



Fonte: A autora (2020)

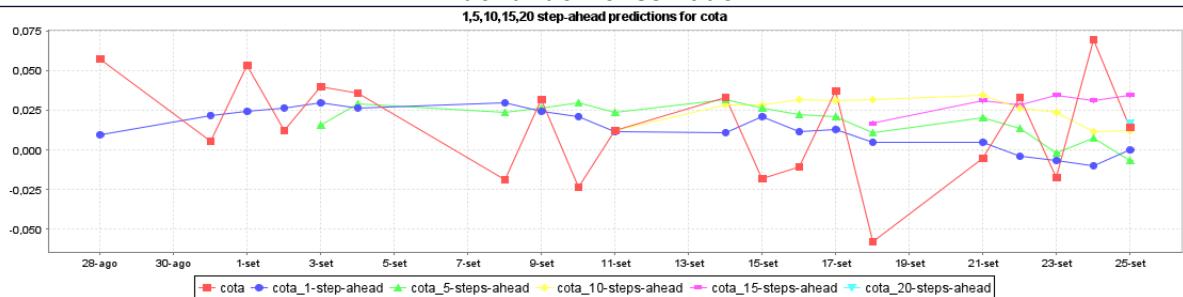
Com a execução do algoritmo para a base completa do fundo conservador (tabela 21), é possível perceber que os erros aumentam conforme o tempo de previsão aumenta, porém na última previsão encontra um resultado melhor que todos os anteriores.

Tabela 21 – Algoritmo de Rede Neural aplicado a base completa do fundo Conservador

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0276	0,0342	0,0012
5 dias futuros	16	0,0283	0,0346	0,0012
10 dias futuros	11	0,0307	0,0413	0,0017
15 dias futuros	6	0,0378	0,0438	0,0019
20 dias futuros	1	0,0023	0,0023	0

Fonte: A autora (2020)

Para esse tipo de base, percebe-se que a reta de previsão está mais distante do dado real (figura 20), se comparado com os gráficos anteriores. Porém, é perceptível como as retas de previsões tem um comportamento parecido, principalmente entre os dias 15 e 25 de setembro.

Figura 20 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Conservador

Fonte: A autora (2020)

b) Fundo Moderado

Da mesma forma em que o fundo moderado foi testado para o algoritmo de Regressão Linear, também foi verificado para a Rede Neural. A tabela 22 apresenta a RNA aplicada a base de um ano do fundo Moderado.

Tabela 22 – Algoritmo de Rede Neural aplicado a base de um ano do fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0800	0,1082	0,0117
5 dias futuros	16	0,0913	0,11	0,0121
10 dias futuros	11	0,0476	0,0575	0,0033
15 dias futuros	6	0,1155	0,1279	0,0163
20 dias futuros	1	0,0379	0,0379	0,0014

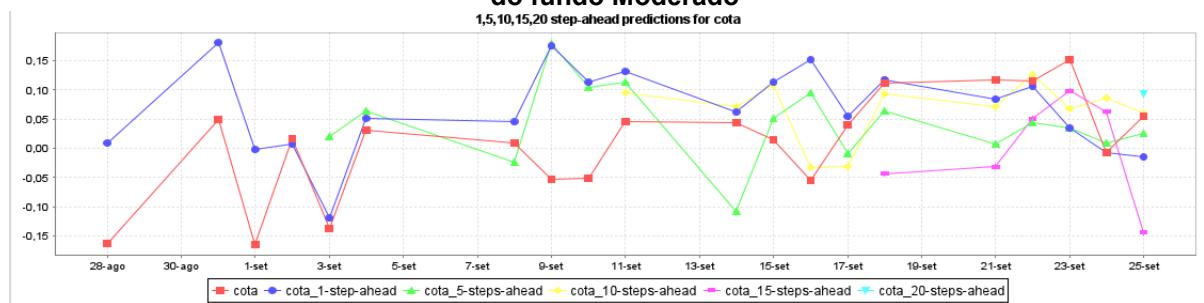
Fonte: A autora (2020)

Para a previsão da base de um ano do fundo moderado, não há relação clara entre os dias futuros e o erro previsto, o que diminui e aumenta conforme se altera o

tempo de previsão. Ainda após essa oscilação, o algoritmo encontra a melhor resultado ao prever o 20º dia.

Pelo gráfico (Figura 21) comparativo gerado, é possível perceber como as previsões estão acompanhando a variação original da cota, com um comportamento parecido ao longo do mês, apesar de oscilar nos valores previstos.

Figura 21 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Moderado



Fonte: A autora (2020)

Para a previsão com a base de dois anos (tabela 23), observa-se uma melhor previsão até chegar no tempo de previsão de 10 dias futuros, com a melhor previsão e posteriormente o erro voltou a aumentar, conforme os dias futuros de previsão aumentaram.

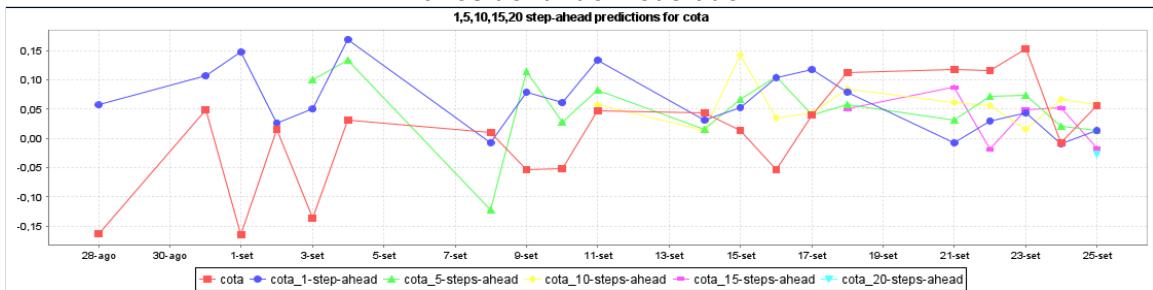
Tabela 23 – Algoritmo de Rede Neural aplicado a base de dois anos do fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0977	0,1247	0,0156
5 dias futuros	16	0,0828	0,1027	0,0105
10 dias futuros	11	0,0563	0,0721	0,0052
15 dias futuros	6	0,0771	0,084	0,0071
20 dias futuros	1	0,0814	0,0814	0,0066

Fonte: A autora (2020)

Diferente dos resultados gerados pela base de um ano, alguns pontos de previsão da base de dois anos (figura 22) não acompanham o comportamento do dado real, como pode-se observar principalmente entre 20 e 17 de setembro.

Figura 22 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Moderado



Fonte: A autora (2020).

Quando o algoritmo é treinado com a base completa (tabela 24), demonstrou-se um comportamento de diminuir os erros previstos conforme os dias futuros de previsão vão aumentando, chegando ao melhor resultado na previsão de 20 dias futuros.

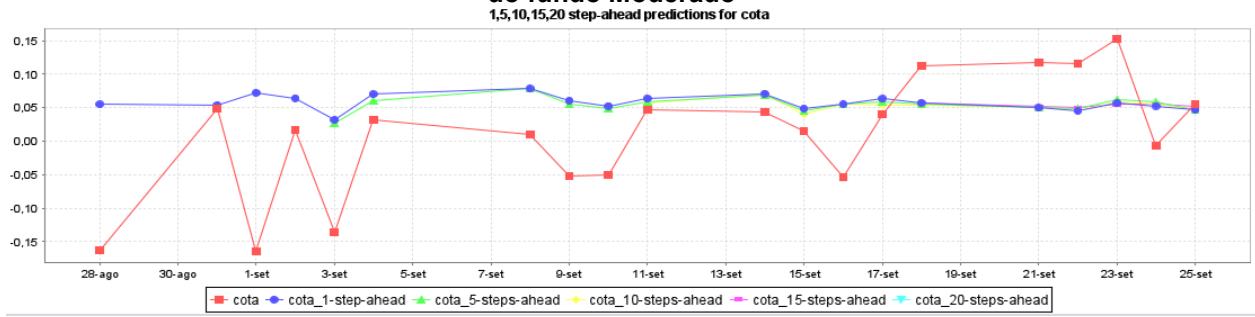
Tabela 24 – Algoritmo de Rede Neural aplicado a base completa do fundo Moderado

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	0,0784	0,1010	0,0102
5 dias futuros	16	0,0639	0,0762	0,0058
10 dias futuros	11	0,0497	0,0598	0,0036
15 dias futuros	6	0,0585	0,0650	0,0042
20 dias futuros	1	0,002	0,0020	0

Fonte: A autora (2020)

Pelo gráfico (Figura 23) gerado é possível perceber claramente com os resultados previstos estão muito próximos entre todos os dias de previsão, ficando em momentos uma curva em cima da outra. Porém, esse tipo de previsão não mostrou claramente o acompanhamento do comportamento como o gráfico gerado pelo algoritmo treinado e testado com a base de um ano.

Figura 23 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Moderado



Fonte: A autora (2020)

c) Fundo Agressivo

E para o último teste da Rede Neural, o algoritmo foi treinado e testado com as bases do fundo Agressivo. A tabela 25 apresenta o RNA para a base de um ano do fundo agressivo.

Tabela 25 – Algoritmo de Rede Neural aplicado a base de um ano do fundo Agressivo

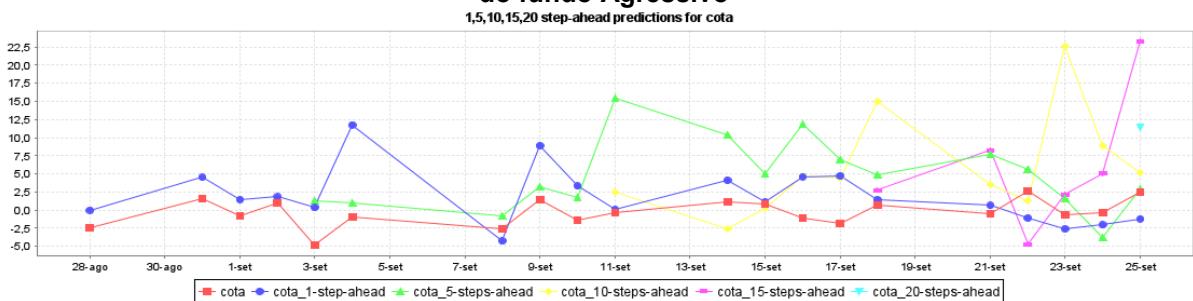
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	3,4449	4,5301	20,5221
5 dias futuros	16	5,4434	6,9204	47,8919
10 dias futuros	11	6,7538	9,3174	86,8144
15 dias futuros	6	7,8755	10,0628	101,2603
20 dias futuros	1	8,9216	8,9216	79,5954

Fonte: A autora (2020)

Pode-se observar que os erros gerados são maiores que os do fundo conservador e moderado, o que já era esperado devido à alta oscilação e volatilidade dos fundos agressivos. Percebe-se que conforme aumenta a previsão de dias futuros, o erro aumenta e o MSE é maior, pelo fato de tratar o erro na forma quadrática, aumentando quando existe algum *outliers*.

Com o gráfico (figura 24) é possível perceber como as previsões com menos dias futuros está mais próxima do dado real, o que acontece com a reta azul (previsão de 1 dia futuro) em relação a vermelha (dados reais).

Figura 24 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de um ano do fundo Agressivo



Fonte: A autora (2020)

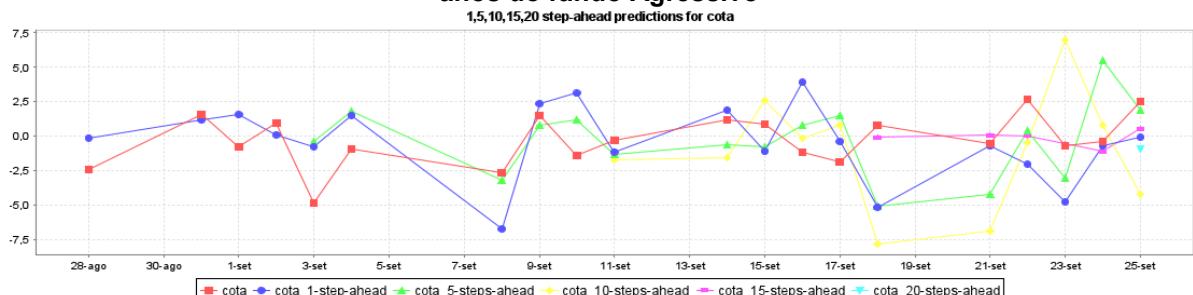
Com a base de treinamento maior, o algoritmo obteve melhor atuação, e diminuiu os erros gerados (tabela 26). Na base de um ano há os maiores erros em 8,92, 10,06 e 101,26 v.s. 3,99, 4,76 e 22,73, em MAE, RMSE e MSE respectivamente.

Tabela 26 – Algoritmo de Rede Neural aplicado a base de dois anos do fundo Agressivo

Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	2,4898	3,0543	9,3285
5 dias futuros	16	2,5999	3,0793	9,4823
10 dias futuros	11	3,922	4,7686	22,7396
15 dias futuros	6	1,1451	1,4338	2,0558
20 dias futuros	1	3,4571	3,4571	11,9518

Fonte: A autora (2020)

Devido à melhor previsão em relação ao teste anterior, é possível verificar pelo gráfico (figura 25) as retas previstas mais próximas da reta real e acompanhando o comportamento da mesma. Apesar de ainda apresentar alguns pontos fora, como por exemplo, o do dia 18 de setembro.

Figura 25 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base de dois anos do fundo Agressivo

Fonte: A autora (2020)

Conforme base de treinamento aumentou, passando de 233 e 484 para 1587 dados, o algoritmo obteve melhor resposta de previsão e assim erros menores em todas as instâncias de previsão (tabela 27).

Tabela 27 – Algoritmo de Rede Neural aplicado a base completa do fundo Agressivo

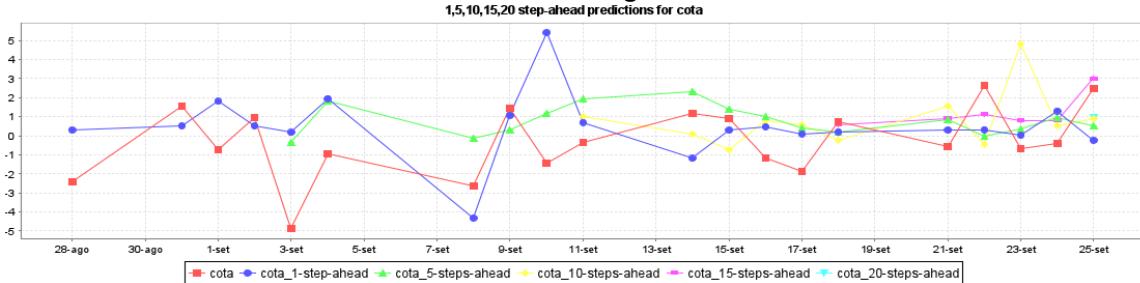
Previsão	N	MAE	RMSE	MSE
1 dia futuro	20	2,0067	2,551	6,5076
5 dias futuros	16	1,9371	2,1744	4,7282
10 dias futuros	11	2,0683	2,4155	5,8347
15 dias futuros	6	1,0638	1,1864	1,4075
20 dias futuros	1	1,5326	1,5326	2,3489

Fonte: A autora (2020)

Com o gráfico gerado (Figura 26) é possível perceber melhor previsão para as previsões futuras, aproximando as retas de previsão com a reta real e seguindo o comportamento da mesma, apesar de ainda ter pontos fora, que pode ser explicado

pelo fato do fundo agressivo ter uma volatilidade maior assim alguns dias a variação da cota é realmente alto.

Figura 26 - Dados reais e previstos pelo algoritmo de Rede Neural, aplicado a base completa do fundo Agressivo



Fonte: A autora (2020)

4.4 Comparação de Desempenho

Para uma visualização de qual algoritmo de se comportou melhor, foram criados os *heatmaps* conforme as Figuras 27, 28 e 29. Desta maneira é demonstrada a diferença do erro de previsão entre a Rede Neural e Regressão Linear: quanto maior for esse número, quer dizer que houve uma melhor resposta pela Regressão Linear e quanto menor for, uma melhor assertividade pela Rede Neural.

Figura 27 – Heatmap do erro MAE para os algoritmos de previsão

a) Fundo Conservador			b) Fundo Moderado			c) Fundo Agressivo		
	1 ano	2 anos		1 ano	2 anos		1 ano	2 anos
1 dia futuro	-0.0112	-0.0021	-0.0057	-0.0197	0.0227	0.0025	1.5106	0.8956
5 dias futuros	-0.0006	-0.0002	0.0002	0.0120	0.0202	0.0052	3.4926	0.8482
10 dias futuros	0.0051	0.0009	-0.0006	-0.0198	0.0022	-0.0015	5.2774	2.5307
15 dias futuros	0.0086	0.0041	0.0058	0.0259	0.0039	-0.0142	6.5089	-0.1456
20 dias futuros	0.0745	0.0007	-0.0047	0.0080	0.0736	-0.0237	7.3311	1.6288
	Base Completa			Base Completa			Base Completa	

Fonte: A autora (2020)

Figura 28 – Heatmap do erro RMSE para os algoritmos de previsão

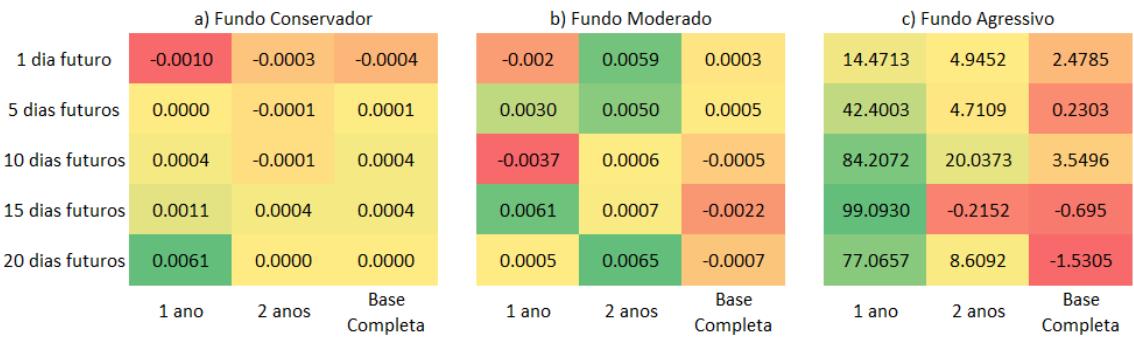
a) Fundo Conservador			b) Fundo Moderado			c) Fundo Agressivo		
	1 ano	2 anos		1 ano	2 anos		1 ano	2 anos
1 dia futuro	-0.0118	-0.0048	-0.0063	-0.0089	0.0262	0.0016	2.0703	0.9607
5 dias futuros	-0.0001	-0.0011	0.0008	0.0146	0.0285	0.0034	4.5770	0.8950
10 dias futuros	0.0060	-0.0012	0.0053	-0.0263	0.0041	-0.0043	7.7027	3.1247
15 dias futuros	0.0135	0.0043	0.0052	0.0268	0.0039	-0.0149	8.5906	-0.0732
20 dias futuros	0.0745	0.0007	-0.0047	0.0080	0.0736	-0.0237	7.3311	1.6288
	Base Completa			Base Completa			Base Completa	

Fonte: A autora (2020)

Considerando os erros MAE e RMSE, observa-se que para o fundo conservador a Regressão Linear se mostrou mais eficiente entre o treinamento com a base de um e dois anos, porém com a base completa a Rede neural apresentou melhores resultados. Já para o fundo moderado, ocorreu comportamento médio entre a base de um e dois anos, mas com a base completa a Rede neural se mostrou mais eficiente, assim como no fundo conservador.

Analizando o fundo agressivo, a Rede Neural se mostrou mais eficiente do que a Regressão Linear com as bases de dois anos e completa, com uma eficiência elevada para a base completa. Por fim, para a base de um ano de treinamento, a regressão linear mostrou melhores resultados.

Figura 29 – Heatmap do erro MSE para os algoritmos de previsão



Fonte: A autora (2020)

Ao ser verificado o erro do MSE pela Figura 29, o comportamento entre o algoritmo de Rede Neural e Regressão Linear se manteve o mesmo, porém com valor de erro superior, principalmente para os treinamentos com os dados do fundo agressivo. Esse maior erro acontece pelo fato de existir algumas variações altas da cota, que podem ser consideradas como *outliers*, potencializando esse erro.

Analizando de modo geral para as várias métricas e diferentes bases de dados nas três categorias de fundos, a RNA se mostrou mais eficaz para as previsões, porém a regressão trouxe também bons resultados, principalmente para a bases de dados com um ano de série histórica, demonstrando como os algoritmos podem ser utilizados para a previsão envolvendo os Fundos de Investimentos.

5. CONCLUSÕES

O presente trabalho teve como objetivo principal prever a rentabilidade futura de fundos de investimentos com o uso de técnicas de *Machine Learning*. Para cumprir tal objetivo, foram investigadas aplicações de *Machine Learning* na área financeira por meio da construção e execução de algoritmos pelo software Weka e, também por meio dos resultados obtidos, foi possível comparar os erros através das tabelas, gráficos e *headmap*, demonstrado que é possível utilizar e aplicar esse tipo de técnica para o mercado financeiro.

Inicialmente foi apresentado referencial teórico sobre investimentos e mineração de dados, com o objetivo de basear e contextualizar a execução e aplicação que se iniciou com teste para parâmetro, execução da Regressão Linear e depois Rede Neural. Como a execução se deu através do Weka (software de interface e interpretação simples), não houve necessidade de pré-conhecimento de programação e linguagens, e foi possível ter a aplicação difundida à investidores leigos quanto à programação.

Com os resultados e erros gerados, foi possível concluir que a aplicação para a previsão de fundos de investimentos é eficaz e pode ser aplicada no cotidiano do mercado financeiro. No contexto da execução formal, como o uso por Assessores de Investimentos para a indicação e construção de uma carteira, é necessária pesquisa e aplicação mais profunda dos métodos, que considere o perfil do algoritmo, com o número de neurônios da camada escondida, números de camada, por exemplo. Também é interessante realizar análises sobre aumento de dias para a previsão futura, trabalhar com mais de dados para treinamentos e testar o comportamento e resultados de diversos fundos, ainda que de mesma categoria.

Porém, para pessoas físicas que já possuem conhecimento prévio sobre o mercado financeiro, é possível utilizar-se do método e aplicar as previsões para seus investimentos atuais e futuros.

Além de ajustes nos algoritmos, manipulação dos dados e alterações referentes aos dados de treinamentos e teste, é importante que no futuro seja avaliada a aplicação por meio de técnicas de redes neurais de aprendizagem profunda (*deep learning*), a quais são muito utilizadas atualmente e podem aprimorar as previsões geradas para os fundos de investimentos.

Também se indica de em vez de utilizar o método de Regressão Linear, testar a aplicação em métodos de auto regressão, como o ARIMA, pois é um tipo de Regressão que avalia apenas os valores passados, podendo obter melhores resultados do que a Linear.

Por fim, entende-se que o tema abordado tem continuidade, com pesquisas e análises que possam dar uma robustez e confiabilidade maior aos modelos e técnicas aplicados. O que se mostra importante para a sociedade e investidores, visto que o mercado financeiro tem ganhado cada vez mais espaço nas mídias e cultura das pessoas, onde a execução pode auxiliar na gestão dos investimentos e aplicações.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, Edgar de. CURSO. **Agente Autônomo de Investimento.** Disponível em: <https://ead.eacertificacoes.com.br/>. Acesso em: 20 maio 2019.

ASSAF, Alexandre. **Mercado Financeiro.** São Paulo: Editora Atlas, 2012.

AMBIMA. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Brasileiro investe pouco e com risco baixo, mostra pesquisa.** Disponível em: http://www.anbima.com.br/pt_br/noticias/brasileiro-investe-pouco-e-com-risco-baixo-mostra-pesquisa.htm. Acesso em: 10 nov. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Menos da metade dos brasileiros têm dinheiro aplicado em produtos financeiros.** Disponível em: http://www.anbima.com.br/pt_br/noticias/menos-da-metade-dos-brasileiros-tem-dinheiro-aplicado-em-produtos-financeiros.htm. Acesso em: 10 nov. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **O raio X do investidor brasileiro.** Disponível em: <http://www.anbima.com.br/data/files/AE/31/E6/CB/52A356107653125678A80AC2/Relatorio-Raio-X-Investidor-PT.pdf>. Acesso em: 10 nov. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Tributação de Ativos e Operações Financeiras – IR e IOF.** Disponível em: http://www.anbima.com.br/data/files/30/53/09/D5/83B61610D6264416A8A80AC2/Mapal_incentivos_v1_2018.pdf. Acesso em: 10 nov. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Nova Classificação dos fundos.** Disponível em: https://www.anbima.com.br/data/files/B4/B2/98/EF/642085106351AF7569A80AC2/Cartilha_da_Nova_Classificacao_de_Fundos_1_.pdf. Acesso em: 15 dez. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Fundos de Investimentos.** Disponível em: http://www.anbima.com.br/pt_br/informar/ranking/fundos-de-investimento/gestores.htm. Acesso em: 15 dez. 2019.

_____. Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. **Principais Custos sobre os investimentos.** Disponível em: <https://comoinvestir.anbima.com.br/escolha/compreensao-do-mercado/principais-custos-sobre-os-investimentos/>. Acesso em: 10 nov. 2019.

B3. Bolsa, Brasil e Balcão. **Debêntures.** Disponível em: http://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-fixa-privada-e-publica/debentures.htm. Acesso em: 10 nov. 2019.

_____. Bolsa, Brasil e Balcão. **Ações.** Disponível em: http://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/acoes.htm. Acesso em: 10 nov. 2019.

BACEN. **Composição e segmentos do Sistema Financeiro Nacional.** Disponível em: <https://www.bcb.gov.br/pre/composicao/composicao.asp>. Acesso em: 10 nov. 2019.

_____. **O que é e o que faz o Banco Central.** Disponível em: <https://www.bcb.gov.br/pre/portalcidadao/bcb/bcfaz.asp?idpai=laipiinstucional>. Acesso em: 10 nov. 2019.

_____. **O Conselho Monetário Nacional.** Disponível em:
<https://www.bcb.gov.br/pre/composicao/cmn.asp>. Acesso em: 10 nov. 2019.

_____. **Fundos de investimentos.** Disponível em:
<http://www.bcb.gov.br/?APLICACOESFAQ>. Acesso em: 10 nov. 2019.

BERRY, Michael; LINOFF, Gordon. **Data mining techniques:** for marketing, sales, and customer relationship management. John Wiley & Sons, 2004.

BERTO, Rosa Maria Villares de Souza; NAKANO, Davi. Revisitando a produção científica nos anais do Encontro Nacional de Engenharia de Produção. **Production**, v. 24, n. 1, p. 225-232, 2014.

BNP PARIBAS. **REGULAMENTO DO WESTERN ASSET FUNDO DE INVESTIMENTO AÇÕES BDR NÍVEL I (“FUNDO”).** Disponível em:
http://www.westernasset.com.br/pt/pdfs/products/6417_regulamento_bylaw.pdf. Acesso em: 10 out. 2020.

BRASIL. Conversão da Medida provisória nº 567, de 2012. **Diário Oficial da União:** Brasília, DF, 2012.

CASTRO, Leandro; FERRARI, Daniel. **Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações.** 6^a ed. Rio de Janeiro: Saraiva, 2016.

CHINO, Daniel Yoshinobu Takada. **Mineração de padrões frequentes em séries temporais para apoio à tomada de decisão em agrometeorologia.** Dissertação (Mestrado em Ciências da Computação e Matemática Computacional) – Universidade de São Paulo, São Paulo, 2014.

CVM. Comissão de Valores Mobiliários. **Sobre a CVM.** Disponível em:
http://www.cvm.gov.br/menu/acesso_informacao/institucional/sobre/cvm.html. Acesso em: 20 nov. 2019.

_____. Comissão de Valores Mobiliários. **Fundos.** Disponível em:
<http://www.cvm.gov.br/menu/regulados/fundos/sobre.html>. Acesso em: 09 dez. 2019.

DAYCOVAL. **RENDA FIXA F.I CRED. PRIV.** Disponível em:
<https://www.daycoval.com.br/Asset/fundos/RendaFixa?lang=ptbr>. Acesso em: 10 out. 2020.

_____. **Histórico.** Disponível em:
<https://www.daycoval.com.br/Asset/asset/historico?lang=ptbr>. Acesso em: 10 out. 2020.

FACELLI, Katti et al. **Inteligência Artificial:** uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC, 2011.

FORTUNA, Eduardo. **Mercado Financeiro, Produtos e Serviços.** 18^a. ed. Rio de Janeiro: Qualitymark, 2010.

GERHARDT, Tatiana; SILVEIRA, Denise. **Métodos de Pesquisa.** 1^a. ed. Porto Alegre: UFRGS, 2009.

GUPTA, Rajta. **Getting started with Neural Network for regression and Tensorflow.** Disponível em: <https://medium.com/@rajatgupta310198/getting-started-with-neural-network-for-regression-and-tensorflow-58ad3bd75223>. Acesso em: 23 nov. 2020.

INFOFUNDOS. **Página Inicial.** Disponível em: <https://infofundos.com.br/>. Acesso em: 30 out. 2019.

_____. **Cotas.** Disponível em: <https://infofundos.com.br/cotas>. Acesso em: 30 set. 2020.

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

GIL, Antônio. **Métodos e técnicas de pesquisa social.** 6ª. ed. São Paulo: Atlas, 2008.

JGP. **Disclaimer.** Disponível em: <https://www.jgp.com.br/quem-somos>. Acesso em: 30 de set. 2020.

_____. **Fundo Perfil.** Disponível em: <https://www.jgp.com.br/fundo-perfil?slug=jgp-strategy-fic-fim>. Acesso em: 20 de set. 2020.

LAROSE, Daniel; CHANTAL, Larose. **Discovering Knowledge in Data: An Introduction to Data Mining.** John Wiley and Sons, Inc, 2005.

LEMES, Antônio; RIGO, Cláudio; CHEROBIM, Ana Paula. **Administração financeira: princípios, fundamentos e práticas brasileiras.** 3ª ed. Rio de Janeiro: Elsevier, 2011.

MORALES-CASTROI, Arturo; RAMÍREZ-REYES, Eliseo; SANABRIA-LANDAZÁBAL, Néstor Juan. MEXICAN STOCK EXCHANGE PERFORMANCE AFTER THE FINANCIAL CRISIS OF 2008: APPLICATION OF DATA MINING. **Dimensión Empresarial**, v. 18, n. 1, p. 28-38, 2020.

MORETTIN, Pedro; TOLOI, Clélia. **Análise de séries temporais.** 2ª. ed. São Paulo: Edgard Blücher, 2006.

_____. BUSSAB, Wilton. **Estatística Básica.** 5ª ed. Rio de Janeiro: Saraiva, 2010.

OLIVEIRA, Mauri Aparecido de; MONTINI, Alessandra de Ávila; BERGMANN, Daniel Reed. Previsão de retornos de ações dos setores Financeiro, de Alimentos, Industrial e de Serviços, por meio de RNA e modelos Arima-Garch. **RAM. Revista de Administração Mackenzie**, v. 9, n. 1, p. 130-156, 2008.

PARK, Jieun; YI, Dokkyun; JI, Sangmin. A Novel Learning Rate Schedule in Optimization for Neural Networks and Its Convergence. **Symmetry**, v. 12, n. 4, p. 660, 2020.

PARTNERS, Fk. CURSO. **Agente Autônomo de Investimento.** Disponível em: <https://fkpartners.com/loja/aaionline/>. Acesso em: 05 out. 2019.

PORTAL DO INVESTIDOR. **O que é uma ação?** Disponível em: https://www.investidor.gov.br/menu/Menu_Investidor/valores_mobiliarios/Acoes/o_que_e_uma_acao.html. Acesso em: nov. 2019.

SANTOS, Bruno. **DM Aula 6 - Tarefas de regressão (ou estimação).** Disponível em: <https://www.youtube.com/watch?v=uvFYhvd1kS>. Acesso em: 05 out. 2020.

SOLIN, Mohammad Maholi et al. Forecasting Portfolio Optimization using Artificial Neural Network and Genetic Algorithm. In: **2019 7th International Conference on Information and Communication Technology (ICoICT)**. IEEE, 2019. p. 1-7.

TERRA, Leonardo Augusto Amaral; PASSADOR, João Luiz. Redes neurais artificiais na previsão da inflação: aplicação como ferramenta de apoio à análise de decisões financeiras em organizações de pequeno porte. **RAM. Revista de Administração Mackenzie**, v. 13, n. 1, p. 68-86, 2012.

VERIOS BLOG. **O que são fundos de investimento.** Disponível em: <https://verios.com.br/blog/fundos-de-investimento-o-que-sao/>. Acesso em: 10 dez. 2019.

VIEIRA, Renato; THOMÉ, Antonio Carlos Gay. Avaliação de redes neurais aplicadas à previsão de índices de mercados de ações. **Relatório Técnico NCE**, 2000.

WANG, Ju-Jie et al. Stock index forecasting based on a hybrid model. **Omega**, v. 40, n. 6, p. 758-766, 2012.

WEKA. **The workbench for machine learning.** Disponível em: <https://www.cs.waikato.ac.nz/ml/weka/>. Acesso em: 20 out. 2020.

WESTERN ASSET. **WESTERN ASSET FIA BDR Nível I.** Disponível em: http://www.westernasset.com.br/pt/pdfs/products/6417_commercialsheet.pdf. Acesso em: 30 set. 2020.

XP Investimentos. **CDB - Certificado de Depósito Bancário.** Disponível em: <https://www.xpi.com.br/investimentos/renda-fixa/cdb/>. Acesso em: 20 nov. 2019.

XP Investimentos. **Renda Fixa.** Disponível em: <https://www.xpi.com.br/investimentos/renda-fixa/lc/>. Acesso em: 20 nov. 2019.

_____. **Fundos de Investimento.** Disponível em: <https://www.xpi.com.br/investimentos/fundos-de-investimento/lista/#/024f0717-bd0f-43ff-8d11-1d75cd54dfd7/detalhes>. Acesso em: 20 nov. 2019.

_____. **O que são Debêntures Incentivadas.** Disponível em: <https://www.youtube.com/watch?v=KDnBEAi-OLE>. Acesso nov. 2019.

YAN, Hongcheng et al. Predicting net asset value of investment fund based on BP neural network. **IEEE**, v. 10, 2010.

ZHONG, Xiao; ENKE, David. Forecasting daily stock market return using dimensionality reduction. **Expert Systems with Applications**, v. 67, p. 126-139, 2017.