

Universidade Federal de Minas Gerais  
Escola de Engenharia  
Curso de Graduação em Engenharia de Controle e Automação

**Predição de Movimentos de Preços no  
Mercado de Ações Brasileiro Usando Modelos  
Baseados em Redes Neurais Artificiais**

Guilherme Magalhães Costa Pinto

Orientador: Prof. Cristiano Leite de Castro

Belo Horizonte, Março de 2021



## **Monografia**

### **Predição de Movimentos de Preços no Mercado de Ações Brasileiro Usando Modelos Baseados em Redes Neurais Artificiais**

Monografia submetida à banca examinadora designada pelo Colegiado Didático do Curso de Graduação em Engenharia de Controle e Automação da Universidade Federal de Minas Gerais, como parte dos requisitos para aprovação na atividade Projeto Final de Curso II.

Belo Horizonte, Março de 2021

# Resumo

O presente trabalho tem como objetivo estruturar um modelo baseado na técnica de redes neurais artificiais, o qual deve prover sinais de compra para ativos financeiros do mercado de capitais brasileiro. Para isso, foram selecionadas 15 empresas listadas na bolsa de valores de São Paulo e 34 indicadores técnicos e fundamentalistas os quais serviram de alimentação para o modelo. Após as definições iniciais, a rede neural foi estruturada e realizou-se, então, predições para um horizonte de até 5 dias. Para cada ativo e dia alvo, foram realizadas 20 simulações, resultando em um total de 1500 simulações. Estas foram avaliadas com base em 4 métricas de desempenho, sendo elas: erro médio quadrático (EMQ), *percentage of change in direction* (POCID), coeficiente de explicação (R<sup>2</sup>) e coeficiente de incerteza (U-Theil). Foi criada uma estratégia de compra e venda dos ativos com base em combinações dos sinais de compra gerados pela rede para cada um dos dias alvo. Por fim, foram realizadas simulações nas quais, estratégias de investimento e estratégias de seleção da rede, que utilizaram do POCID como balizador para a escolha, foram combinadas. Em uma das simulações o modelo alcançou um resultado de 73% de rentabilidade, superando todos os *benchmarks* estabelecidos com pelo menos 80% de superioridade. O resultado obtido reforça o propósito da pesquisa de conseguir identificar padrões de comportamento nas cotações dos ativos e dessa forma, utilizar dados passados para realizar a previsões de preços do mercado acionário brasileiro.



# Abstract

The present work aims to structure a model based on the technique of artificial neural networks, which should provide buy signals for financial assets in the Brazilian capital market. For this, 15 companies listed on the São Paulo stock exchange and 34 technical and fundamentalist indicators were selected, which served as input for the model. After the initial definitions, the neural network was structured and predictions were made for the next 5 days. For each stock and target day, 20 simulations were carried out, resulting in a total of 1500 simulations. These were evaluated based on 4 performance metrics, namely: mean squared error (EMQ), percentage of change in direction (POCID), explanation coefficient (R2) and uncertainty coefficient (U-Theil). A strategy for buying and selling stocks was created based on combinations of the buy signals generated by the network for each target day. Finally, simulations were carried out in which investment strategies and network selection strategies, which used POCID as a marker, were combined. In one of the simulations, the model achieved a result of 73% profitability, surpassing all the benchmarks established with at least 80% superiority. The result obtained reinforces the purpose of the research to be able to identify patterns of behavior in the quotations of the stocks and to use past data to carry out the price forecasts of the Brazilian stock market.



# Agradecimentos

À Universidade Federal de Minas Gerais (UFMG), por prover todos os recursos necessários para minha formação.

Ao Prof. Cristiano Leite de Castro, pela orientação, me apoiando com prontidão nos momentos em que precisei.

Aos professores do curso de Engenharia de Controle e Automação, por todos os ensinamentos e incentivos.

Ao colegas de curso, em especial ao amigo Guilherme Cramer, por toda ajuda e momentos divididos durante todo o curso.

Aos amigos da Consultoria e Projetos Elétricos Júnior (CPE), por dividirem momentos de incertezas e superações.

À minha namorada Leticia Murta, por todo incentivo nos momentos de angústia e pela revisão gramatical.

Em especial a minha família que sempre me deu todo o suporte durante minha trajetória acadêmica.





# Sumário

<b>Resumo</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Agradecimentos</b>	<b>v</b>
<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>Lista de Siglas</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos do Projeto . . . . .	2
1.2 Estrutura da Monografia . . . . .	3
<b>2 Fundamentação Teórica</b>	<b>5</b>
2.1 Mercado Financeiro . . . . .	5
2.1.1 Análise Técnica . . . . .	6
2.1.2 Indicadores Técnicos . . . . .	7
2.1.3 Análise Fundamentalista . . . . .	8
2.1.4 Indicadores Fundamentalistas . . . . .	8
2.2 Redes Neurais . . . . .	9
2.2.1 Função de ativação . . . . .	11
2.2.2 Overfitting e Underfitting . . . . .	11
2.2.3 Arquitetura de uma RNA . . . . .	13
2.2.4 Rede Neural Recorrente - RNN . . . . .	14
2.2.5 Long Short Term Memory - LSTM . . . . .	15
2.3 Resumo do Capítulo . . . . .	17

<b>3</b>	<b>Metodologia</b>	<b>19</b>
3.1	Dados . . . . .	19
3.1.1	Entradas e saídas do modelo . . . . .	21
3.1.2	Análise de correlação . . . . .	24
3.1.3	Normalização dos dados . . . . .	25
3.2	Seleção da Rede Neural . . . . .	26
3.2.1	Configuração da rede . . . . .	26
3.2.2	Medida de Desempenho . . . . .	27
3.3	Modelo de compra e venda . . . . .	28
3.4	Resumo do Capítulo . . . . .	29
<b>4</b>	<b>Resultados</b>	<b>31</b>
4.1	Requisitos do Sistema . . . . .	31
4.2	Desenvolvimeto e Implementação . . . . .	32
4.3	Testes . . . . .	33
4.3.1	Simulação 1 . . . . .	33
4.3.2	Simulação 2 . . . . .	33
4.3.3	Simulação 3 . . . . .	35
4.3.4	Simulação 4 . . . . .	36
4.4	Análise dos Resultados . . . . .	38
4.5	Desempenho Comparativo . . . . .	40
<b>5</b>	<b>Conclusões</b>	<b>45</b>
5.1	Considerações Finais . . . . .	46
5.2	Propostas de Continuidade . . . . .	46
	<b>Referências Bibliográficas</b>	<b>48</b>

# Lista de Figuras

2.1	Exemplo de regressão. Os pontos representam as entradas $x$ e $y$ e a linha representa a função estimada no intervalo $[0$ a $10]$ <i>Fonte:</i> Svetlana Cheusheva . . . . .	10
2.2	Exemplo de classificação. Amostras de duas classes distintas, indicadas por cores diferentes e separadas pela reta verde. <i>Fonte:</i> Saumya Saxena . . . . .	10
2.3	Exemplo de previsão. Os pontos azuis representam as entradas $x$ e $y$ e o ponto vermelho representa a previsão do próximo valor da sequência <i>Fonte:</i> Wanessa Pereira dos Reis . . . . .	10
2.4	Quatro principais funções de ativação. Função degrau, função sigmoidal, função linear e função hiperbólica. <i>Fonte:</i> Prof Eduardo Simas . . . . .	12
2.5	Exemplo de overfitting, onde a rede tenta acertar todos os pontos. <i>Fonte:</i> Edpresso Team . . . . .	12
2.6	Exemplo de underfitting, onde a rede não consegue identificar uma relação entre as variáveis de entrada. <i>Fonte:</i> Edpresso Team . . . . .	13
2.7	Exemplo de saída desejada. <i>Fonte:</i> Edpresso Team . . . . .	13
2.8	Rede neural artificial multicamadas. Cada parâmetro da entrada ( $X$ ) serve de entrada para todos os neurônios da primeira camada escondida e assim sucessivamente até que a camada de saída produza ( $Y$ ). <i>Fonte:</i> Anderson Castro Soares de Oliveira . . . . .	14
2.9	Neurônio de uma rede neural recorrente. A seta em loop indica a realimentação da informação para ser utilizado no próximo processamento. <i>Fonte:</i> Autor . . . . .	15
2.10	Diferença entre uma RNN padrão e uma LSTM. Na figura 1 temos uma RNN padrão com a aplicação da não linearidade tanh. A figura 2 mostra uma LSTM com a aplicação de 4 não linearidades. <i>Fonte:</i> Christopher Olah . . . . .	16
3.1	Divisão dos dados em grupos de treino, teste e simulação. <i>Fonte:</i> Autor . . . . .	20

3.2	Cenários para ativação do sinal de compra. Caso as três previsões de um cenário forem de valorização, a compra será executada na abertura do dia posterior. <i>Fonte: Autor</i> . . . . .	29
4.1	Rentabilidades dos banchmarks comparados à simulação 4. <i>Fonte: Autor</i>	41
4.2	Rentabilidades dos banchmarks comparados à simulação 4 ao longo do tempo. <i>Fonte: Autor</i> . . . . .	42
4.3	Rentabilidades dos fundos de ações e do modelo de redes neurais. <i>Fonte: Autor</i> . . . . .	43

# Lista de Tabelas

3.1	Ativos - Empresas analisadas e seus respectivos códigos . . . . .	20
3.2	Entradas do modelo - Indicadores que servirão de entrada para o modelo	21
3.3	Coeficiente de correlação entre a cotação de abertura de BBDC4 e as outras variáveis . . . . .	25
3.4	Arranjos de redes testados . . . . .	26
4.1	Médias dos resultado das simulações de predição para “d+1” . . . . .	32
4.2	Médias dos resultado da fase de testes para todo o horizonte de predição	33
4.3	Médias dos resultado da fase de simulação para todo o horizonte de predição . . . . .	33
4.4	Resultados da simulação 1 . . . . .	34
4.5	Resultados da simulação 2 . . . . .	34
4.6	Sinais de compra. . . . .	35
4.7	Retornos. . . . .	36
4.8	Resultados da simulação 3 . . . . .	37
4.9	Resultados da simulação 4 . . . . .	37
4.10	Sinais de compra. . . . .	38
4.11	Retornos. . . . .	39
4.12	Resumo das configurações e resultados das simulações . . . . .	39
4.13	Resumo das configurações e resultados das simulações . . . . .	41



# Lista de Siglas

<i>B3</i>	Brasil, Bolsa, Balcão
<i>CDI</i>	Certificado de Depósito Interbancário
<i>DFC</i>	Demonstrativo de Fluxo de Caixa
<i>DRE</i>	Demonstração de Resultados do Exercício
<i>EMQ</i>	Erro Médio Quadrático
<i>IFR</i>	Índice de Força Relativa
<i>IMAT</i>	Índice de Materiais Básicos
<i>IMOB</i>	Índice Imobiliário
<i>INDX</i>	Índice do Setor Industrial
<i>IPCA</i>	Índice de Preços no Consumidor
<i>LSTM</i>	Long Short Term Memory
<i>MACD</i>	Média Móvel Convergente Divergente
<i>MME</i>	Média Móvel Exponencial
<i>PIB</i>	Produto Interno Bruto
<i>POCID</i>	Percentage of Change Direction
<i>RNA</i>	Rede Neural Artificial
<i>RNN</i>	Rede Neural Recorrente
<i>RP</i>	Regressão Polinomial



# Capítulo 1

## Introdução

A bolsa de valores brasileira vive sua fase mais promissora, atingindo recordes de investidores a cada dia. Porém, ainda existe a ideia de que os investimentos são baseados em apostas e seus retornos dependem exclusivamente da sorte do investidor. Indo além desse pensamento a teoria do *random walk* [1] defende que as variações nos valores das ações independem uma da outra. Com isso, uma análise que se baseia nos movimentos passados no mercado financeiro não é eficiente.

Na contramão do pensamento popular, existem teorias as quais defendem que as séries financeiras têm memória de longo prazo, ou seja, os preços de um ativo acionário seguem padrões de comportamento. Em [2] o autor propõe uma análise baseada no coeficiente de Hurst, o qual, por meio da estatística, busca verificar se há um padrão de memória ligado a um evento ou se este não passa de um acontecimento aleatório. A solução proposta fundamenta-se na ideia de que as séries financeiras aqui abordadas seguem padrões de comportamento. Portanto torna-se válida a tentativa de descobrir, por meios computacionais, o padrão de repetição ligado a uma determinada série financeira.

Dessa forma, modelos baseados Redes Neurais Artificiais se tornam atrativos para a resolução do problema, uma vez que “RNAs se apresentam como uma das possíveis e mais populares abordagens para a resolução de problemas de RP. Não obstante, as RNAs se aplicam a uma gama maior de problemas e de áreas do conhecimento, não estando a mesma restrita a problemas de RP. Por serem aproximadores universais de funções [Cyb89] as RNAs podem ser aplicadas também a problemas de regressão e de previsão.” [3]. Outro ponto importante é a não linearidade dos dados, visto que o mercado financeiro é volátil e passa por colapsos e crescimentos repentinos, conhecidos como bolhas especulativas. Portanto, reforça-se a escolha pelo uso de um modelo baseado em RNAs, os quais lidam com dados não-lineares.

Por fim, é necessário cautela ao avaliar o desempenho do modelo criado em termos

de precisão da previsão como mostrado em [4] *“choosing the wrong accuracy metric can be very misleading when evaluating model performance.”*, isso ocorre pois os dados de uma série temporal tendem a ser correlacionados no tempo, ou seja, o índice no momento “ $t+1$ ” provavelmente estará próximo do índice no momento “ $t$ ”, logo *“what the model is actually doing is that when predicting the value at time “ $t+1$ ”, it simply uses the value at time “ $t$ ” as its prediction”* [4].

Um fator importante para alcançar bons resultados é a qualidade dos dados de entrada, uma vez que o algoritmo deverá reconhecer padrões de comportamento na série histórica para prever o preço futuro do ativo. Frente a isso optou-se por utilizar indicadores técnicos e fundamentalistas como base de dados para alimentar o modelo, sendo esses melhor descritos nas seções 2.1.2 e 2.1.4.

Em função disso, o presente trabalho busca responder o seguinte questionamento: é possível construir um modelo baseado na técnica de redes neurais artificiais capaz de utilizar dados passados para prever preços de ativos do mercado acionário brasileiro, com o objetivo de alcançar rentabilidades atrativas e acima da média para o investidor?

## 1.1 Objetivos do Projeto

Objetivo geral: criar um modelo baseado em RNAs para fazer predições de preços de ações

Objetivos específicos

- a) Coletar e tratar os dados;
- b) Construir o modelo;
- c) Verificar a eficiência desse modelo;
- d) Criar uma estratégia eficiente de compra e venda dos ativos;
- e) Comparar o modelo com outros investimentos e com as carteiras criadas como *benchmarks*;

Por último, espera-se que o trabalho contribua para a área de finanças quantitativas, em especial, para a utilização de redes neurais artificiais na predição de séries temporais financeiras e ao uso de dados passados como base para as predições. Dessa forma busca-se estimular o aprimoramento e desenvolvimento de novos projetos e técnicas que utilizam sistemas computacionais inteligentes voltadas para aplicações no mercado financeiro.

## 1.2 Estrutura da Monografia

O trabalho está dividido em cinco capítulos. Este capítulo apresentou uma introdução ao projeto desenvolvido. O Capítulo 2 descreve os princípios básicos do mercado financeiro e das RNAs, além de abranger todos os conceitos necessários para um melhor entendimento do projeto. O Capítulo 3 aborda a metodologia de desenvolvimento, seguido pelo capítulo 4 onde a implementação do modelo e as validações são apresentadas. No Capítulo 5 tem-se a conclusão da monografia e algumas sugestões e dificuldades encontradas na realização do projeto.



# Capítulo 2

## Fundamentação Teórica

Neste capítulo buscou-se apresentar e descrever todos os conhecimentos utilizados para a realização do presente trabalho. Nas seções seguintes serão abordados conceitos de contabilidade, mercado financeiro, análise técnica e seus indicadores, os quais serão fundamentais para a definição de quais dados alimentarão o modelo. Além disso, serão descritos os conceitos relacionados ao aprendizado de máquina e redes neurais artificiais, possibilitando assim, a escolha da melhor configuração para o modelo.

### 2.1 Mercado Financeiro

Mercado financeiro é a denominação dada ao ambiente onde são negociados os mais diversos produtos financeiros, dentre eles ações, *commodities* e câmbio monetário. De forma organizada e bem estruturada, compradores e vendedores são colocados num mesmo local - atualmente virtual, através das corretoras de valores - onde podem comprar e vender um ativo financeiro. Intuitivamente é possível notar que tal mercado, como qualquer outro, é gerido pela lei da oferta e da procura, portanto o valor de uma ação, por exemplo, tende a aumentar caso o número de interessados em comprá-la seja maior que a quantidade ofertada.

Usualmente os ativos são negociados em bolsas de valores, as quais centralizam todas as operações realizadas. De acordo com [5], a primeira bolsa de valores com as características modernas que se tem registro surgiu na Bélgica no ano de 1487, local onde comerciantes se reuniam para negociar a compra e venda de metais preciosos, moedas e câmbio. Somente a partir do século XIX o comércio de ações se instaurou, elevando o volume de transações nas bolsas de valores bem como sua especialização separando as bolsas de mercadorias das bolsas de valores imobiliários.

No Brasil temos a BMFBOVESPA, denominada de B3 (Brasil, Bolsa, Balcão), onde estão centralizadas todas as operações do mercado financeiro brasileiro. Foi criada em

2000, a partir da unificação das bolsas de São Paulo (BOVESPA), do Rio de Janeiro (BVRJ), de Minas-Espírito Santo-Brasília (BOVMESB), do Extremo Sul (BVES), de Santos, da Bahia-Sergipe-Alagoas (BOVESBA), de Pernambuco e Paraíba (BOVAPP), do Paraná (BVPR) e a Regional (BVRg). No ano de 2020 o montante negociado na bolsa brasileira atingiu a marca de R\$6,45 trilhões valor igual a 93,1% do PIB projetado para o ano [6].

### 2.1.1 Análise Técnica

Análise técnica, ou análise gráfica de ações é a ferramenta utilizada por investidores para analisar o movimento do preço de um ativo financeiro, tendo como objetivo principal identificar tendências de alta ou baixa e assim comprar ou vender o ativo. A análise técnica moderna se baseia na Teoria de Dow, formulada por Charles Henry Dow em 1884. Em seu trabalho, Charles descreve as tendências do mercado e como elas se comportam, baseando-se em 6 princípios.

- **O mercado leva em consideração tudo:** os preços refletem todos os fatores que afetam a oferta e procura do mercado
- **O mercado se movimenta com base em tendências:** primária, a qual reflete o longo prazo durando meses ou anos, secundária, que tem duração de 1/3 a 2/3 da primária e terciária responsável pelas correções e ajustes dos preços em alguns dias ou semanas.
- **Foco na tendência primária:** essa é composta por 3 fases, acumulativa, onde alguns investidores realizam suas posições, participação, onde mais alguns investidores notam a tendência e fazem suas posições e por último vem a fase de realização, onde a tendência é confirmada.
- **Confirmação:** os índices mais abregentes do mercado, como o Dow Jones devem seguir e confirmar a tendência.
- **Volume de vendas:** para que a tendência mude é preciso um aumento expressivo no volume de negociações
- **Tendências se substituem:** uma tendência permanecerá até que um fator externo a altere.

Nesse sentido, percebe-se a complexidade do mercado e como a procura e identificação de tendências é importante para a obtenção de bons resultados.

### 2.1.2 Indicadores Técnicos

Existem uma série de indicadores utilizados na análise técnica que auxiliam no estudo do comportamento do mercado. Estes, em sua maioria, são derivados dos preços do ativos e carregam particularidades que, combinados, auxiliam na análise de uma série financeira. De modo geral os indicadores técnicos podem ser enquadrados em 5 categorias, de acordo com seu objetivo e de como é calculado.

- **Médias Móveis:** É o indicador mais conhecido e amplamente usado. Sua ideia é simples, utiliza-se a média de um período anterior para antecipar um valor futuro. Seu uso é ideal para séries com forte tendência, porém podem apontar falsos sinais quando a tendência não é bem definida.
- **Tendências:** Esse tipo de indicador, como o próprio nome diz, tenta mostrar quando há uma tendência na série financeira. Uma técnica utilizada para a identificação de tendência é o MACD (Convergência / Divergência das Médias Móveis), esse indicador combina 3 MMEs (Média Móvel Exponencial) para tentar sinalizar a mudança de tendência em uma série. Seu cálculo é feito da seguinte maneira:

Linha MACD = (MME 12 dias – MME 26 dias)

Linha de Sinal = MME 9 dias

Histograma MACD = Linha MACD – Linha de Sinal

- **Momento:** O momento varia de acordo com a variação do preço de um ativo. Quanto mais rápido for o aumento do preço, maior será o momento. A estabilização do preço gera uma gradual diminuição de momento. Observar esse indicador permite ao investidor determinar quando está ocorrendo uma situação de sobre compra ou sobre venda do ativo.
- **Volatilidade:** A volatilidade busca medir o desvio da medida em relação a uma média, portanto quanto maior a volatilidade, maior é a oscilação de preço de um determinado ativo. Esse indicador ajuda o investidor a determinar o que esperar no comportamento dos preços de um ativo.
- **Volume:** Indicadores de volume usam o volume de operações e a frequência com que os ativos são comprados ou vendidos. Olhar para esse indicador permite definir se os movimentos nos preços são realmente relevantes ou são fruto de operações pontuais.

### 2.1.3 Análise Fundamentalista

Análise fundamentalista baseia-se na observação e interpretação da situação econômica e financeira de uma determinada empresa, setor, moeda, país e as expectativas e projeções geradas em torno de seus indicadores. Em suma esse tipo de análise acontece com base nos demonstrativos de resultados e balancetes gerados periodicamente pelas empresas e governo - quando a análise é feita a nível macroeconômico - tendo em vista que dados econômicos nacionais e internacionais podem influenciar o desempenho das empresas.

Imagine por exemplo uma empresa que tem como principal atividade a extração, refinamento e venda de petróleo e seus derivados. Para o investidor é importante saber qual o volume de vendas do trimestre passado, quanto disso foi investido na busca de novos poços de extração de petróleo, quais os tamanhos das reservas, quais são os impostos pagos por essa empresa, qual o preço pago pelo barril de petróleo no exterior dentre inúmeras outras informações as quais influenciam na tomada de decisão de tornar-se sócio ou não da companhia.

Tendo em vista todos os aspectos envolvidos na análise fundamentalista, cabe ao investidor tentar identificar o potencial de crescimento do lucro de uma empresa no futuro, algo que, geralmente, leva à valorização do ativo. Para tomar tal decisão são levadas em conta três classes de indicadores, as quais serão mostradas e discutidas na próxima sessão.

### 2.1.4 Indicadores Fundamentalistas

Assim como na análise técnica, na análise fundamentalista existem indicadores responsáveis por demonstrar e explicar a situação econômica e financeira de uma empresa. Estes se dividem em três principais grupos:

- **Indicadores Macroeconômicos:** Tratam-se de medidas do cenário macroeconômico em que a empresa está inserida. Em geral são ligados ao ambiente onde os negócios dessa empresa ocorrem e seu níveis de atividade econômica. Normalmente é a etapa inicial da análise fundamentalista, nela deve-se considerar dados como o PIB (produto interno bruto), taxas de câmbio e de juros, nível de desemprego, entre outros.
- **Indicadores Setoriais:** Além dos agregados macroeconômicos é preciso refinar e atentar-se a alguns desses indicadores, pois podem afetar isoladamente setores diferentes. Um exemplo são os incentivos fiscais concedidos pelo governo a determinadas áreas, além da questão política que altera regulamentações e podem



gerar até mudanças estruturais.

- **Indicadores da Empresa:** Por fim deve-se olhar para os dados financeiros da empresa com o objetivo final de determinar o valor intrínseco, ou preço justo das ações. Esses dados normalmente advem do balanço patrimonial, da DRE (demonstração de resultados do exercício) e do DFC (demonstrativo do fluxo de caixa).

Em suma uma análise fundamentalista correta dá ao investidor a possibilidade de projetar o comportamento futuro dos preços de negociação de uma empresa com base em números presentes. Essa visão aparentemente simples esconde inúmeros outros fatores os quais podem influenciar o mercado financeiro. Apenas a análise fundamentalista ou técnica são ferramentas insuficientes para lidar com o objetivo da previsão do preço de um ativo seja para amanhã ou para daqui a seis meses.

Com o passar do tempo cada vez mais informação é necessária para se compreender os problemas do mercado financeiro, tornando essa tarefa ainda mais complexa para o ser humano, sendo portanto, necessária a presença de máquinas capazes de processar enormes quantidades de dados e mais que isso, gerar alguma informação útil a partir desse processamento.

## 2.2 Redes Neurais

Redes neurais artificiais são modelos computacionais inteligentes baseados no sistema nervoso central humano, mais precisamente em neurônios, que são capazes de adquirir conhecimento através da experiência. Uma rede neural trata basicamente de 3 tipos de problemas:

- **Problemas de Regressão:** a solução para esse problema consiste em estimar uma função aproximadora a qual demonstra a relação entre variáveis a partir de amostras de dados. Segundo [3], “a capacidade de aproximação não-linear com um número de parâmetros que cresce apenas linearmente como o número de variáveis”, portanto RNAs se mostram interessantes, como métodos de resolução de problemas que envolvem múltiplas variáveis. Um exemplo de regressão pode ser visto na figura 2.1.
- **Problemas de Classificação:** nesse caso a solução consiste na associação de amostras às classes pré determinadas. Para isso, uma função discriminante é utilizada para realizar partições no espaço de entradas e assim determinar as classes. Um exemplo de classificação pode ser visto na figura 2.2.

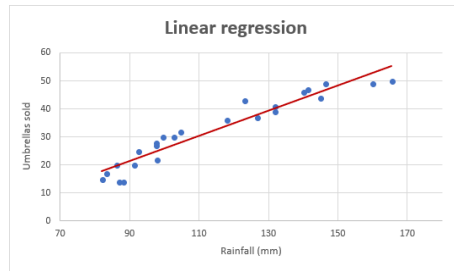


Figura 2.1: Exemplo de regressão. Os pontos representam as entradas  $x$  e  $y$  e a linha representa a função estimada no intervalo  $[0 \text{ a } 10]$  *Fonte: Svetlana Cheusheva*

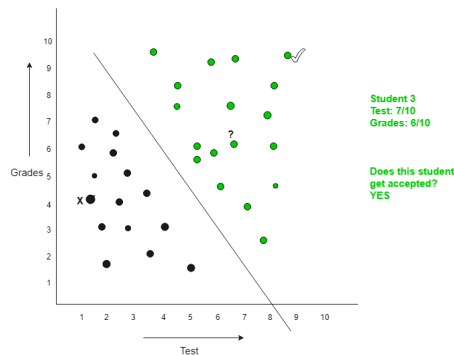


Figura 2.2: Exemplo de classificação. Amostras de duas classes distintas, indicadas por cores diferentes e separadas pela reta verde. *Fonte: Saumya Saxena*

- Problemas de Previsão:** esse é o problema o qual o presente trabalho se propõe a resolver, de forma similar ao problema de regressão é preciso estimar a relação entre variáveis, porém, neste caso há uma relação temporal entre elas. Um ponto importante a ser levado em consideração é um possível atraso de tempo em que uma variável pode influenciar no alvo da previsão, esta é denominado *lag*. Segundo [3] sua identificação pode ser feita através de uma análise de correlação entre as variáveis ou utilizando o conhecimento prévio sobre o problema. Um exemplo de previsão pode ser visto na figura 2.3.

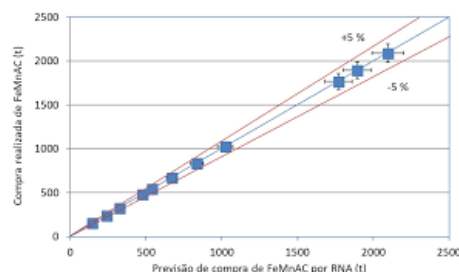


Figura 2.3: Exemplo de previsão. Os pontos azuis representam as entradas  $x$  e  $y$  e o ponto vermelho representa a previsão do próximo valor da sequência *Fonte: Wanessa Pereira dos Reis*

As RNAs são formadas, tal como o cérebro humano, por uma interconexão de unidades de processamento, essas denominadas neurônios. Os neurônios são responsáveis por executar operações matemáticas e têm a capacidade de armazenar o conhecimento adquirido a partir dessas operações e torná-lo disponível para uso. Para que isso seja possível, é necessária uma atribuição de pesos entre as conexões dos neurônios, dessa forma, pode-se controlar o impacto de cada variável de entrada na saída do modelo.

Uma das principais características das redes neurais artificiais é a capacidade de aprendizado, isso só é possível pois, durante a fase de treinamento do modelo, os pesos atribuídos à saída de cada neurônio são ajustados com base na medida de desempenho da rede.

### 2.2.1 Função de ativação

A função de ativação de uma rede neural tem um papel crucial na resposta do modelo, pois é ela a responsável por determinar se um neurônio será ativado ou não, ou seja, se a informação será propagada à entrada do próximo neurônio ou saída da rede ou se será ignorada. O formato da função de ativação varia de acordo com a necessidade da rede e o modelo escolhido, mas tipicamente é representada pela equação:

$$\mu = \sum_{i=1}^N (w_i * x_i) = W^T X ,$$

em que  $x_i$  são os elementos do vetor de entrada  $X$  e  $w_i$  são os pesos correspondentes advindos do vetor de pesos  $W$  [3]. Existem inúmeras funções de ativação, dentre as quais algumas são mostradas na Figura 2.4

### 2.2.2 Overfitting e Underfitting

O overfitting ocorre quando a rede se adapta muito bem aos dados durante o processo de treinamento de forma que a saída é sempre muito próxima da realidade, ou seja, a rede basicamente está decorando o comportamento de uma amostra dos dados, porém na fase de testes a rede tentará reproduzir esse comportamento para uma amostra diferente dos dados gerando assim um desempenho ruim.

O underfitting o modelo não consegue encontrar a relação entre as variáveis, tornando o desempenho ruim já na fase de treinamento. Geralmente esse problema ocorre quando é utilizado um número menor de neurônios que o necessário na arquitetura da rede.

Deve-se tomar cuidado redobrado na avaliação de desempenho do modelo, pois é possível que as métricas não identifiquem tais problemas. A figura 2.7 mostra a saída

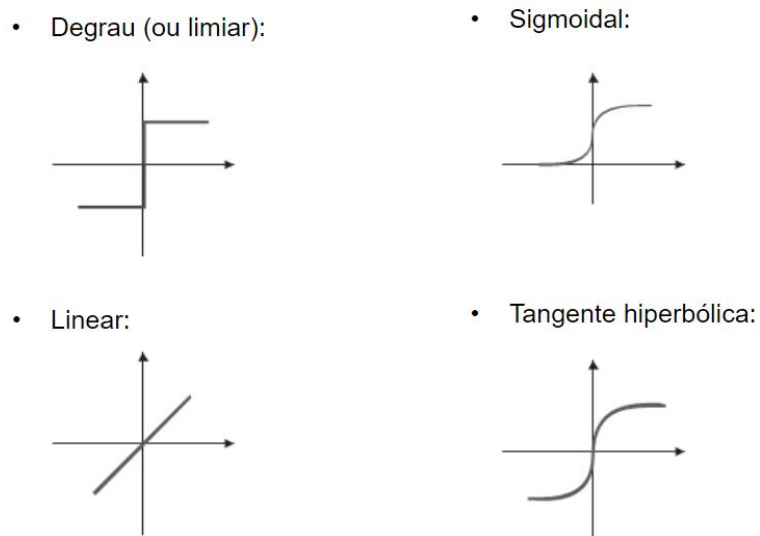


Figura 2.4: Quatro principais funções de ativação. Função degrau, função sigmoidal, função linear e função hiperbólica. *Fonte:* Prof Eduardo Simas

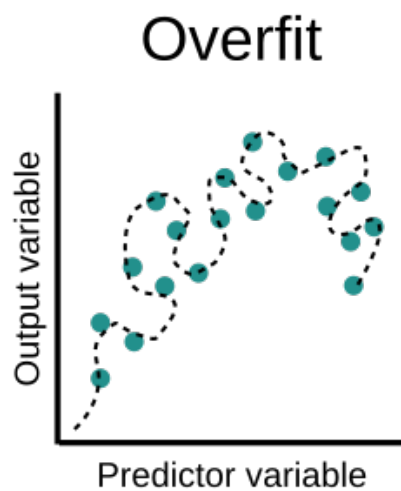


Figura 2.5: Exemplo de overfitting, onde a rede tenta acertar todos os pontos. *Fonte:* Edpresso Team

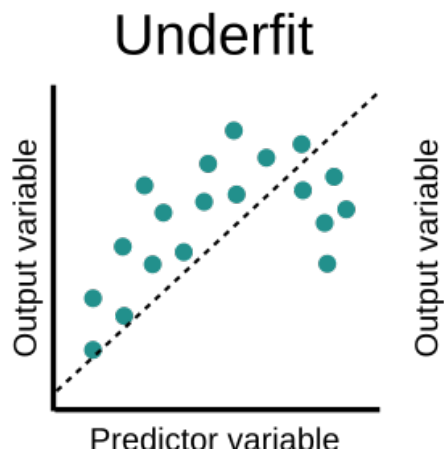


Figura 2.6: Exemplo de underfitting, onde a rede não consegue identificar uma relação entre as variáveis de entrada. *Fonte:* Edpresso Team

desejada para o exemplo:

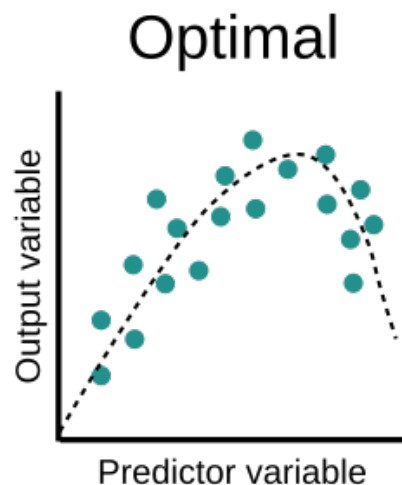


Figura 2.7: Exemplo de saída desejada. *Fonte:* Edpresso Team

### 2.2.3 Arquitetura de uma RNA

Como dito anteriormente, uma RNA é formada por interconexões entre neurônios, portanto a maneira como essas unidades são agrupadas definem o tipo de arquitetura da rede. A estrutura e o funcionamento básico de uma rede neural é apresentado a seguir:

- Os sinais de entrada são passados para cada um dos neurônios.
- Cada um desses sinais é multiplicado por um peso, que indica a influência dessa entrada para a saída do neurônio.

- Uma soma ponderada é feita com os produtos de entrada-peso gerando um nível de atividade para cada neurônio.
- Caso este nível ultrapasse um determinado limite, o neurônio produz um sinal de saída, o qual pode ser a resposta final da rede ou simplesmente a entrada para uma próxima camada de neurônios.

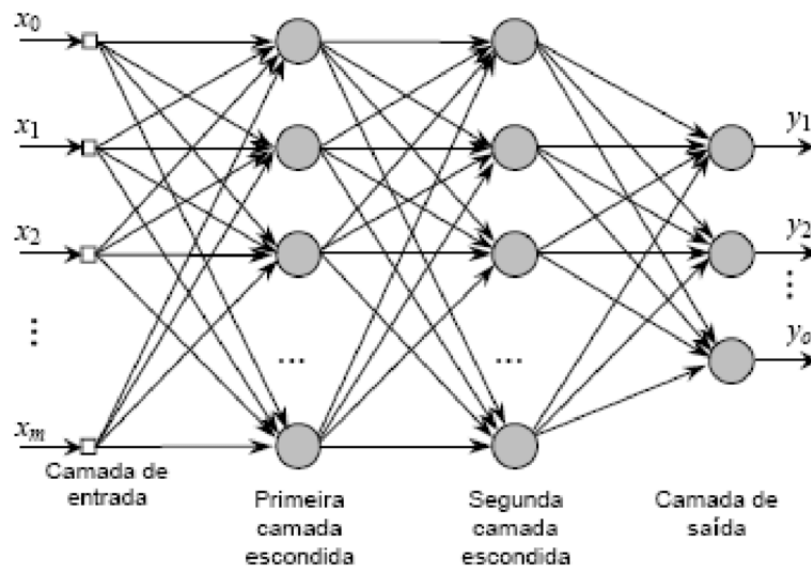


Figura 2.8: Rede neural artificial multicamadas. Cada parâmetro da entrada (X) serve de entrada para todos os neurônios da primeira camada escondida e assim sucessivamente até que a camada de saída produza (Y). *Fonte:* Anderson Castro Soares de Oliveira

A definição da melhor arquitetura é um dos grandes desafios na resolução de um problema, uma vez que esse é um processo experimental de tentativa e erro e demanda tempo de execução. Na prática o processo consiste em testar diferentes métodos de aprendizagem combinado às infinitas possibilidades de configurações da rede - número de camadas e quantos neurônios farão parte de cada uma delas [7].

## 2.2.4 Rede Neural Recorrente - RNN

Uma classe de problemas relacionando as redes neurais artificiais requer maior cautela na sua avaliação e solução, pois envolvem a variável temporal, não só o tempo em si, mas também uma relação de dependência sequencial. As redes neurais tradicionais têm deficiências para processar esses tipos de dados, já que cada resposta do modelo depende apenas das entradas e dos pesos, ou seja, não há correlação temporal entre os dados.

RNNs resolvem esse problema pois possuem a habilidade de persistir a informação referente ao processamento anterior, ou seja, as entradas anteriores afetam o resultado da próxima saída e assim sucessivamente. Portanto, além dos pesos, cada neurônio é capaz de armazenar um estado referente a entrada em “t-1”.

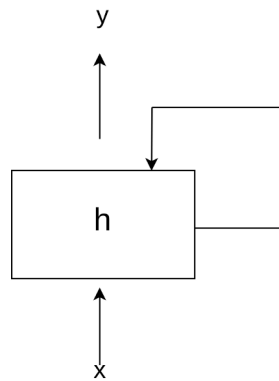


Figura 2.9: Neurônio de uma rede neural recorrente. A seta em loop indica a realimentação da informação para ser utilizado no próximo processamento. *Fonte:* Autor

Um exemplo prático de aplicação de uma RNN é a previsão de palavras em um determinado texto. Para se prever a próxima palavra em uma frase é necessário o conhecimento prévio das palavras anteriores, além disso, a ordem em que elas estão dispostas também influenciam no resultado.

### 2.2.5 Long Short Term Memory - LSTM

Um exemplo de RNNs são as LSTMs, essas redes são capazes de aprender dependências de longo prazo, ou seja, persistem as informações de tempo por mais de um ciclo, permitindo então formar contextos de longo prazo. Essa característica permitirá o desenvolvimento do presente trabalho, haja visto que a predição de uma série do mercado financeiro depende de múltiplos fatores com múltiplas relações temporais.

As LSTMs são redes formadas com o intuito de resolver o problema da dependência de longo prazo, portanto seu comportamento padrão é lembrar-se de informações por longos períodos de tempo, além é claro, da característica básica de uma rede recorrente que é utilizar o estado anterior para realizar o processamento.

RNNs padrão têm forma de uma cadeia de módulos repetidos, onde a saída “h-1” alimenta o próximo módulo e assim sucessivamente. A estrutura interna é simples, formada por apenas uma não linearidade “tanh” aplicada a combinação linear entre a matriz de pesos pela entrada  $X_t$  e os dados da camada oculta no tempo anterior  $h_{t-1}$ . Já os LSTMs têm a estrutura interna diferente, pois contemplam 4 não linearidades

interaginado de uma maneira especial a fim de preservar também a memória de longo prazo.

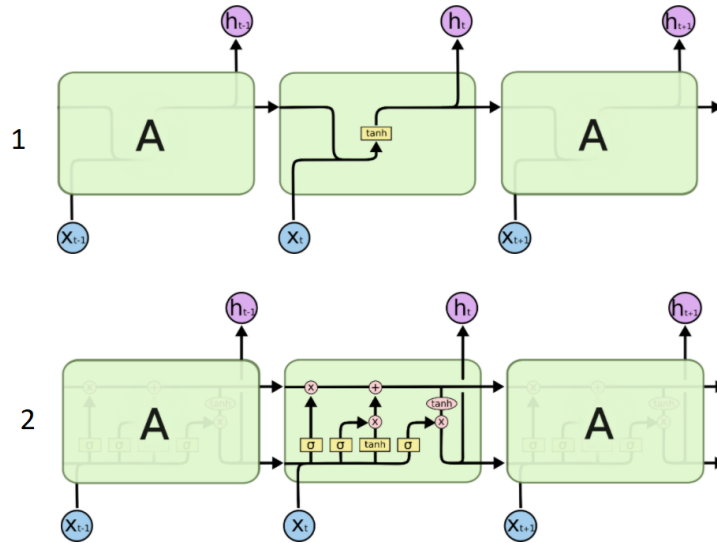


Figura 2.10: Diferença entre uma RNN padrão e uma LSTM. Na figura 1 temos uma RNN padrão com a aplicação da não linearidade tanh. A figura 2 mostra uma LSTM com a aplicação de 4 não linearidades. *Fonte:* Christopher Olah

A ideia central de uma LSTM é o chamado estado da célula, linha horizontal que atravessa o topo do diagrama. Seus dados fluem pela cadeia passando apenas por iterações lineares, ou seja a informação pode facilmente passar inalterada, mantendo-se ativa por muitos ciclos. A atualização e inserção de novas informações é feita em 2 etapas. Primeiramente tem-se a combinação linear entre a matriz de pesos pela entrada  $X_t$  e o estado anterior  $h_{t-1}$ , a qual passará por uma função de ativação do tipo sigmoidal, o resultado, por sua vez, será multiplicado ao estado da célula. Tal operação tem como objetivo forçar a retirada ou conservação de determinada informação do estado da célula.

O segundo passo é adicionar novas informações úteis ao estado da célula. Isso é feito por meio da combinação entre um vetor de novos candidatos, representado pela combinação linear entre  $X_t$ ,  $h_{t-1}$  pela matriz de pesos  $W$  aplicada a uma função de ativação do tipo “tanh” e a mesma combinação linear anterior aplicada a função de ativação sigmoidal. O resultado dessa operação é um vetor de novos candidatos com valores dimensionados de acordo com o quanto serão atualizados cada valor de estado.

Nesse ponto temos um novo vetor  $C_t$ , o qual teve as informações importantes mantidas, as não relevantes removidas e as novas, adicionadas [8][9][10]. Vê-se então a explicação para o nome LSTM, onde L (*long*) é representado pelas informações mantidas por um longo período de tempo e S (*short*) é representado pelas novas informações



adicionadas a cada ciclo.

## 2.3 Resumo do Capítulo

A partir do estudo e detalhamento dos conceitos é possível partir para a fase de planejamento da execução do trabalho buscando entender as técnicas e explorar os recursos referentes aos temas abordados até aqui. Portanto o próximo passo é a elaboração da metodologia, para que assim, seja possível realizar os testes e analisar o desempenho do modelo baseado na técnica de redes neurais artificiais para a predição de preços no mercado acionário brasileiro.



# Capítulo 3

## Metodologia

Para esse estudo utilizou-se uma análise quantitativa na pesquisa do tipo *ex-post-facto*, uma vez que o conjunto de dados utilizados são referentes a acontecimentos já ocorridos e portanto não controláveis. Trata-se também de uma pesquisa bibliográfica, haja visto que foram utilizados materiais de acesso público, como dados de balanço financeiro, cotações e estatísticas de empresas, bem como artigos, livros e notícias do mercado financeiro.

A natureza quantitativa facilita o entendimento dos conceitos e corrobora com a confiabilidade dos resultados, uma vez que análises estatísticas são aplicadas aos dados desde o pré-processamento até o resultado final, tendo assim resultados mais fundamentados.

### 3.1 Dados

O conjunto de dados utilizados foi composto por empresas de capital aberto na Bolsa de Valores de São Paulo no período de 05/10/2017 a 29/01/2021. A análise não se restringiu apenas a empresas que fazem parte do índice Bovespa (IBOVESPA), haja visto que estas possuem grande liquidez, algo que poderia adicionar certo ruído às cotações, portanto decidiu-se também por analisar ativos de baixa liquidez a fim de que o portfólio fosse o mais heterogêneo possível. Dessa forma, foram selecionadas 15 empresas, as quais estão listadas na tabela 3.1.

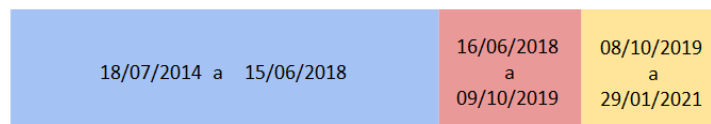
Os dados necessários para a realização da pesquisa foram obtidos por meio do software Economática. Vale ressaltar que as cotações diárias foram ajustadas de forma a considerar pagamentos de dividendos e bonificações, além de considerar possíveis desdobramentos e agrupamentos dos ativos. Para dados faltantes decidiu-se, por simplicidade, por repetir o valor do dia anterior. Seria possível também optar pela realização de uma interpolação ou até mesmo escolher apenas períodos onde não houvessem

Código	Empresa
BBDC4	Bradesco
BRFS3	BRF
BRKM5	Braskem
CSNA3	Siderúrgica Nacional
ELET3	Eletrobras
EMBR3	Embraer
GGBR4	Gerdau
HAPV3	Hapvida
ITUB4	Itaú Unibanco
KLBN4	Klabin
MGLU3	Magazine Luiza
MULT3	Multiplan
OIBR4	Oi
PETR4	Petrobras
VALE3	Vale

Tabela 3.1: Ativos - Empresas analisadas e seus respectivos códigos

faltas, porém esse número é irrelevante diante da massa de dados, não interferindo na análise e cálculos do modelo.

A janela de estudo foi escolhida considerando todo o volume de dados capturados para que houvessem amostras suficientes para o aprendizado do modelo e, ao mesmo tempo, não se utilizassem dados desatualizados, visando assim compreender a dinâmica atual da variação de preços dos ativos. A massa de dados foi separada em 3 partições com 60%, 20% e 20% do total dos dados dando origem à partições de treino, teste e simulação, respectivamente. A imagem 3.1 representa essa separação com base nas datas.

Figura 3.1: Divisão dos dados em grupos de treino, teste e simulação. *Fonte:* Autor

Para esse estudo optou-se por capturar dados diários já para variáveis amostradas com outras frequências foi necessário realizar uma regressão linear simples. Um exemplo é a taxa Selic anualizada, a qual tem suas atualizações feitas mensalmente, nesse caso foi realizada a interpolação entre meses consecutivos.

Indicadores Gerais	Indicadores Específicos
Produção de Óleo	Cotação de fechamento
Empregos formais	Cotação de abertura
Saldo Depósito na poupança	Cotação máxima
Índice de commodities	Cotação mínima
Índice de confiança do consumidor	Quantidade de negociações
CDI	Volume negociado
Dólar	P/L
IBOVESPA	EV/EBITDA
NASDAQ	Dividend Yield 1 mês
S&P500	Dividend Yield 1 ano
Rentabilidade da poupança	Média mensal da cotação
IPCA	Liquidez 1 mês
INDX	Banda de Bollinger 20/2
IMAT	Banda de Bollinger 10/2
IMOB	IFR (9 períodos)
	MMCD 10/30
	MMCD 2/6
	Momento 10 dias
	Momento 3 dias

Tabela 3.2: Entradas do modelo - Indicadores que servirão de entrada para o modelo

### 3.1.1 Entradas e saídas do modelo

Após a escolha dos ativos a serem estudados foi necessário definir os indicadores, os quais servirão de entrada para a rede neural. Esses indicadores, ou *features*, foram divididos em duas classes. A primeira é composta por indicadores gerais, isto é, comuns a todas as empresas e a segunda classe é formada por indicadores específicos, os quais contém informações diretamente relacionados ao ativo. A tabela 3.2 mostra os indicadores com base nessa classificação.

A seguir é apresentada um breve explicação sobre cada um dos indicadores.

- **Produção de Óleo:** Produção de derivados de petróleo.
- **Empregos formais:** Número de pessoas em empregos formais.
- **Saldo depósito na poupança:** Saldo da poupança em milhões de reais.
- **Índice de commodities:** Reflete o desempenho do mercado de commodities.
- **Índice de confiança do consumidor:** Mostra as expectativas atuais e futuras de consumidores e de empresários da indústria em relação à situação econômica.

- **CDI:** É a taxa que lastreia as operações interbancárias e regula o mercado de renda fixa.
- **Dólar:** Cotação do dólar.
- **IBOVESPA:** Cotação do IBOVESPA, índice composto pelas maiores empresas do Brasil.
- **NASDAQ:** Cotação da NASDAQ, índice composto por todas as empresas listadas.
- **S&P500:** Cotação do S&P500, índice composto pelas 500 maiores empresas dos estados unidos.
- **Rentabilidade da Poupança:** Rentabilidade mensal da poupança.
- **IPCA:** É o Índice de Preços para o Consumidor Amplo que busca identificar a variação dos preços no comércio.
- **INDX:** Índice do setor industrial. Tem como objetivo de medir o desempenho das ações mais representativas do setor industrial.
- **IMAT:** Índice de materiais básicos. Indicador de desempenho dos ativos de maior negociabilidade e representatividade do setor de materiais básicos.
- **IMOB:** Índice Imobiliário. Indicador de desempenho da atividade imobiliária compreendida por exploração de imóveis e construção civil.
- **Cotação de fechamento:** Preço de fechamento do período para um ativo.
- **Cotação de abertura:** Preço de abertura do período para um ativo.
- **Cotação máxima:** Preço máximo atingido pelo ativo no período.
- **Cotação mínima:** Preço mínimo atingido pelo ativo no período.
- **Quantidade de negociações:** Quantidade de compras e vendas de um ativo realizadas no período.
- **Volume negociado:** Quantidade de negociações multiplicada pela média do montante das negociações.
- **P/L:** Preço dividido pelo lucro da empresa. Apresenta uma idéia de quanto o mercado está disposto a pagar pela empresa.

- **EV/EBITDA:** É a soma entre o valor de mercado das ações de uma empresa e sua dívida líquida dividido pelo resultado operacional da empresa. Mostra quanto tempo levaria para o resultado operacional pagar o investimento.
- **Dividend Yield 1 mês:** É a quantidade de proventos distribuídos no último mês dividido pela cotação do ativo.
- **Dividend Yield 1 ano:** É a quantidade de proventos distribuídos no último ano dividido pela cotação do ativo.
- **Média da cotação:** Média da cotação de fechamento no período.
- **Liquidez 1 mês:** Média da liquidez do ativo no mês.
- **Banda de Bollinger 20/2:** Indicador de volatilidade que mostra se um ativo está sobre-comprado ou sobre-vendido, ou seja, se está acima ou abaixo do valor ideal. O número 20/2 indica que a média móvel utilizada para seu cálculo é de 20 períodos com mais ou menos 2 vezes o desvio padrão.
- **Banda de Bollinger 10/2:** É a mesma Banda de Bollinger, porém a média móvel é de 10 períodos.
- **IFR (9 períodos):** Esse indicador é uma espécie de termômetro do cansaço do mercado, ou seja, mostra em uma escala de 0 a 100 se o ativo já se valorizou ou desvalorizou muito dentro do período estipulado. Nesse trabalho o IFR é com 9 dias.
- **MMCD 10/30:** É a diferença entre duas médias móveis exponenciais (10 dias e 30 dias). É um indicador de tendência na variação do preço.
- **MMCD 2/6:** É a mesma MMCD, porém a diferença é entre as médias móveis exponenciais de 2 dias e 6 dias.
- **Momento 10 períodos:** O Momento ou Momentum é um indicador de tendência, mostrando com qual rapidez o preço do ativo se valoriza ou desvaloriza em relação ao preço do ativo no período estipulado, nesse caso, 10 dias.
- **Momento 3 períodos:** Mesma medida de Momento, porém em relação aos últimos 3 dias.

Na sequência, os alvos do modelo foram definidos como as cotações de abertura dos ativos para “d+1”, “d+2”, “d+3”, “d+4”, “d+5”, ou seja, a tarefa do modelo seria prever o preço de abertura do ativo para os 5 dias subsequentes ao dia atual. Sendo

assim, “d+1” o dia subsequente, “d+2” dois dias subsequentes e assim por diante. Note que a compra de um ativo só poderá ser realizada na abertura do próximo dia, uma vez que a cotação de fechamento do dia atual também faz parte das entradas do modelo, portanto a predição “d+1” não gerará nenhum tipo de retorno. No entanto optou-se por manter essa predição para que, a partir das métricas de desempenho, fosse mais um balisador na escolha da rede neural a ser utilizada.

A primeira vista presume-se que quanto maior o número de variáveis de entrada, melhor será o resultado do modelo, porém isso nem sempre é verdade. Dentre as *features* selecionadas podem existir ocorrências de variáveis que não tem relação alguma com o alvo, portanto são dados que não contribuem para a resolução do problema e ainda adicionam complexidade desnecessária ao modelo. Dito isso é necessária uma forma de avaliar qual variável realmente está relacionada com o alvo desejado da rede neural.

### 3.1.2 Análise de correlação

Muitas vezes, em problemas que envolvem um grande número de variáveis, é preciso entender como estas se relacionam e como uma interfere nas outras. Um método matemático muito conhecido e efetivo para entender de que forma acontecem esses relacionamentos é a análise de correlação. Para [11] as técnicas baseadas na análise de correlação são ferramentas importantes para diversas áreas, como Ciências Sociais e do comportamento, da Engenharia e das Ciências Naturais.

Como mostrado na seção 3.1.1 foram selecionadas 34 variáveis, com base no estudo prévio de mercado financeiro, dentre elas indicadores técnicos e fundamentalistas e, a partir daí, é necessária uma pré-análise para avaliar a real contribuição de cada uma dessas variáveis para o modelo.

Para analisar e definir quais indicadores influenciam na variável alvo, foi utilizada uma função de correlação entre o preço de abertura do ativo e cada uma das *features* de entrada. Essa métrica mostra em uma escala de 0 a 1 a existência de uma relação estatística entre duas variáveis. Para exemplificar são mostradas na tabela 3.3 as correlações entre o preço de abertura de BBDC4 e cada uma das possíveis entradas do modelo.

Note que o P/L, os Dividend Yield de 1 mês e 1 ano, a Liquidez de 1 mês, o IRF, as Médias Móveis CD 10/30 e 2/6, os Momentos de 10 e 3 dias, o índice de confiança do consumidor, assim como o IMOB possuem correlação menor que 0,4, portanto decidiu-se por não utilizá-los como entradas do sistema, visto que poderiam interferir no resultado. A rentabilidade da poupança, o número de empregos formais e o CDI possuem correlação negativa, porém seu módulo é maior que 0,4, logo também foram utilizadas como entradas do modelo.



Indicador	Correlação	Indicador	Correlação
Produção - Óleo	0.755	Cotação de fechamento	0.9991
Empregos formais	-0.493	Cotação máxima	0.9995
Saldo poupança	0.743	Cotação mínima	0.9995
Índice de commodities	0.648	Qnt. de negociações	0.514
Índice de conf. do consumidor	-0.122	Volume negociado	0.632
CDI	-0.615	P/L	0.075
Dólar	0.718	EV/EBITDA	-
IBOVESPA	0.862	Dividend Yield 1 mês	0.065
NASDAQ	0.842	Dividend Yield 1 ano	0.238
S&P500	0.878	Média cotação	0.993
Poupança acumulado	0.854	Liquidez 1 mês	0.203
Rentabilidade da poupança	-0.435	Banda de Bollinger 20/2	0.993
IPCA	0.905	Banda de Bollinger 10/2	0.997
IMAT	0.734	IFR (9 períodos)	0.070
IMOB	0.388	MMCD 10/30	0.148
INDX	0.875	MMCD 2/6	0.043
		Momento 10 períodos	0.078
		Momento 3 períodos	0.026

Tabela 3.3: Coeficiente de correlação entre a cotação de abertura de BBDC4 e as outras variáveis

Tal operação foi realizada para todos os ativos aqui estudados, de forma que as entradas do modelo de predição fossem compostas apenas por variáveis correlatas ao alvo, ou seja, aquelas as quais o módulo da correlação fosse maior que 0,4.

### 3.1.3 Normalização dos dados

Após a coleta dos dados foi realizada uma análise e estudo das variáveis do problema, pode-se notar que cada uma retrata uma grandeza diferente e por consequência é expressa com base em uma escala específica. Ao alimentar a rede neural com essas variáveis, aquelas com ordem de grandeza superior, acabam por se sobressair às demais, mascarando sua influência na predição da variável alvo.

Para contornar esse problema faz-se necessário realizar a normalização dos dados, processo o qual busca refazer as escalas de todas as variáveis, tornando-a única, com máximo e mínimo pré estabelecidos. [12]. Para isso utilizou-se da função “MinMaxScaler” oriunda da biblioteca “scikit-learn”, a qual provê recursos de aprendizado de máquina para a linguagem de programação Python.

Decidiu-se por normalizar os dados em uma escala de 0 a 1, ou seja, o máximo valor presente na massa de dados assumiu o valor 1 enquanto o menor valor assumiu o valor 0. Os demais valores foram definidos conforme a equação:

Arranjo	Neurônios
1	32 x 16
2	32 x 64
3	128 x 64
4	64 x 32 x 16
5	64 x 64 x 32
6	128 x 64 x 32
7	128 x 128 x 64
8	256 x 128 x 64
9	256 x 256 x 64

Tabela 3.4: Arranjos de redes testados

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}},$$

onde  $X_{min}$  é o menor valor da massa de dados,  $X_{max}$  o maior valor,  $X$  é o valor original e  $X_{scaled}$  é o valor normalizado.

Por fim, tal transformação pode ser revertida, pois os valores das predições também serão feitos com base na normalização, para isso utilizou-se da função *inverse\_transform*, recurso oriundo também da biblioteca “scikit-learn”.

## 3.2 Seleção da Rede Neural

Uma rede neural pode ser composta por infinitas configurações e o primeiro passo é definir qual tipo de rede utilizar. Para esse estudo, optou-se por utilizar uma rede do tipo LSTM, classe de redes as quais apresentam ótimos resultados envolvendo problemas de previsões com séries temporais. As decisões tomadas a partir desse ponto são diretamente influenciadas pelo nível de conhecimento do autor na área da pesquisa, pois esse fator contribui e facilita o alcance de resultado desejado. Portanto, pode-se concluir que apesar das inúmeras possibilidades, cabe ao pesquisador a pré seleção das configurações do modelo baseado na rede LSTM.

### 3.2.1 Configuração da rede

A escolha da configuração de uma rede neural é obtida com base em um trabalho incessante de tentativa e erro, portanto para atingir o objetivo foram pré selecionados 9 possíveis arranjos (tabela 3.4). Cada uma dessas configurações foi testada para 5 ativos selecionados aleatoriamente e realizando as predições de todos os 5 dias alvo.

### 3.2.2 Medida de Desempenho

Na seção anterior foram apresentados os possíveis arranjos para a rede neural, esses foram baseados na metodologia LSTM, discutida na seção 2.2.5. Para as redes 4 a 9 foram adicionados dropouts as camadas 2 e 3, esses são responsáveis por eliminar parte da informação, haja visto que as redes propostas possuem uma grande quantidade de neurônios em suas camadas. Tal processo busca forçar a rede a não se especializar nos dados na fase de treinamento, para que não haja o *overfitting* 2.2.2. Os dropouts utilizados foram de 0,3, ou seja, 30% da informação foi eliminada após a segunda e terceira camadas.

Após a realização das predições é necessário avaliar os resultados, para isso utilizaram-se 4 métricas para entender o quão boas eram as predições feitas pelo modelo. As medidas de desempenho utilizadas foram R2 (coeficiente de explicação), EMQ (erro médio quadrático), POCID (*percentage of change direction*) e U-Theil (coeficiente de incerteza).

O R2 é uma medida de ajuste do modelo aos dados reais, ou seja, expressa em termos de porcentagem, o quão explicativo é o modelo. Como uma medida em forma de porcentagem varia de 0 a 100%, quanto mais próximo de 100%, melhor as predições se ajustam à amostra de dados.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} ,$$

onde  $y_i$  é o valor desejado,  $\hat{y}_i$  é o valor previsto e  $\bar{y}$  é a média dos valores desejados.

O EMQ tem como principal objetivo encontrar a diferença média entre a série de valores desejados e a série de valores preditos. De forma prática, seu objetivo é entender os erros da predição

$$EMQ = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 ,$$

O cálculo do POCID é de extrema importancia na identificação de tendências, uma vez que ele indica os acertos do modelo quanto as previsões de valorizações e desvalorizações das cotações. Este não deve ser analisado separadamente, uma vez que não é uma medida de exatidão.

$$POCID = \frac{1}{n} \sum_{i=1}^n a_i ,$$

$$a_i = \begin{cases} 1 & \text{se } (y_i - y_{i-1})x(\hat{y}_i - \hat{y}_{i-1}) > 0 \\ 0 & \text{se } (y_i - y_{i-1})x(\hat{y}_i - \hat{y}_{i-1}) < 0 \end{cases}$$

O U-Theil é útil para mostrar o quanto os resultados estão bons, comparando a previsão do modelo com a previsão ingênua que diz que a melhor predição é a repetição do valor anterior, ou seja, na previsão ingênua o melhor valor para o preço de amanhã é o preço de hoje.

$$U = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \hat{y}_{i-1})^2}},$$

### 3.3 Modelo de compra e venda

Um dos principais desafios do trabalho é definir como será feita a aplicação do capital, haja visto que essa etapa pode comprometer diretamente o resultado final caso uma decisão equivocada seja tomada. A estratégia como um todo pode ser analisada em duas frentes, quanto a escolha de quando comprar e quando vender e quanto a forma como o capital será investido.

A definição de quando comprar se baseia na saída da rede neural, caso esse valor seja maior que o valor atual da cotação de abertura, existe aí uma indicação de valorização do ativo, ou seja, esse é um sinal de compra. É válido observar que como as previsões são feitas para o preço de abertura, levando em consideração todos os dados do presente dia, o investimento inicial só poderá ocorrer em “d+1”, ou seja, essa previsão já não trará retorno, uma vez que a compra só pode ser exercida no momento para o qual a previsão foi feita.

A estratégia então, foi analisar o horizonte de predição mais distante, a partir de “d+2” até “d+5”. Portanto o sinal de compra seria confirmado após, pelo menos 3 previsões de valorização. Para exemplificar o processo imagine a seguinte situação: em 05/05/2020 uma previsão “d+5” é feita para 10/05/2020, no dia 06/05/2020, uma previsão “d+4” é feita para 10/05/2020 e no dia 07/05/2020 uma previsão “d+3” é feita para 10/05/2020. Caso as 3 previsões indiquem uma valorização relativa para o dia 10 um sinal de compra é gerado, sendo assim no dia 08/05/2020 o ativo é adquirido pelo preço de abertura. As vendas serão feitas na data estipulada da previsão, mesmo que haja previsão de valorização para o próximo dia, no exemplo acima seria executada a venda na abertura do dia 10.

Esse é apenas 1 dos 4 possíveis cenários para que haja um sinal de compra. O diagrama da imagem 3.2, retrata de forma visual todos os cenários.

A segunda decisão importante para completar a estratégia de aplicação do capital é quanto ao montante disponível e qual parcela desse valor seria aplicada caso houvesse um sinal de compra. Primeiramente foi definido o montante inicial disponível para a aplicação, nesse processo o valor escolhido foi de R\$10.000,00, o qual foi definido

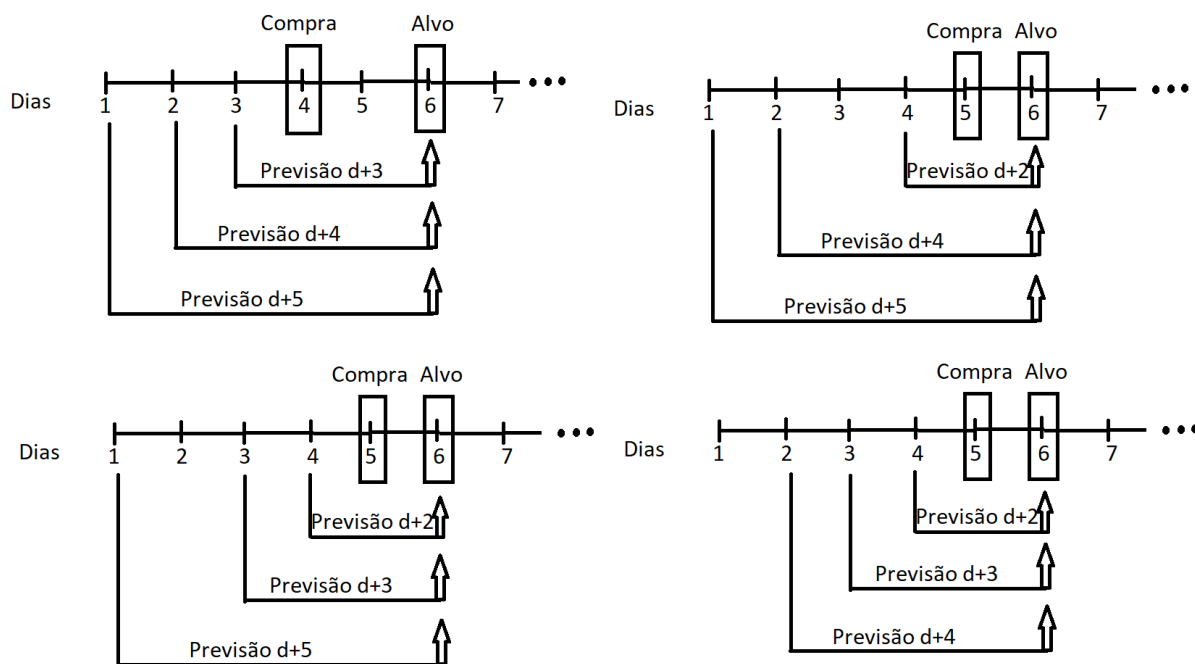


Figura 3.2: Cenários para ativação do sinal de compra. Caso as três previsões de um cenário forem de valorização, a compra será executada na abertura do dia posterior.  
*Fonte:* Autor

aleatoriamente.

Duas estratégias foram selecionadas para que fossem avaliados os retornos e assim escolher aquela que se ajustasse melhor ao modelo proposto. A estratégia 1 definiu que o total investido diariamente seria de R\$5.000,00. Caso não houvessem sinais de compra para nenhum ativo, não haveria investimento naquele dia, caso mais de um ativo tivesse um sinal de compra para aquele dia o montante disponível para aplicação seria dividido igualmente entre eles e se somente 1 ativo tivesse sinal de compra naquele dia, todo o valor seria investido nele. Por fim definiu-se que todas as vendas deveriam ser feitas ao atingir a data alvo das previsões, independente do retorno ser positivo ou negativo. A segunda estratégia foi similar à primeira, porém todo o montante acumulado estaria disponível para aplicação.

## 3.4 Resumo do Capítulo

Nesse capítulo foram apresentadas as estratégias adotadas durante todo o desenvolvimento do trabalho, desde a coleta dos dados até a definição de como se dariam os investimentos. Essa parte do trabalho tem como objetivo apresentar e explicar o porquê de cada decisão e por consequência, fornecer a base para as fases de testes e simulações.

O primeiro passo foi definir quais seriam as empresas estudadas. Nesse sentido

foram escolhidas 15 as quais possuem tamanho, área de atuação, relação com o exterior e características de mercado distintas umas das outras, a fim de tornar o portfólio heterogêneo e remover qualquer viés sazonal ou setorial às cotações.

Posteriormente foram escolhidos os períodos de estudo, levando em consideração a quantidade de dados disponível, fator crucial para que a rede neural conseguisse identificar padrões de comportamentos e tomando o cuidado para trabalhar com dados mais atuais, de forma que a dinâmica atual do mercado fosse identificada.

A escolha do alvo do modelo e de quais variáveis serviriam de alimentação para a rede neural é parte crucial do processo, portanto uma análise de correlação foi necessária para que somente variáveis relacionadas com o alvo fossem consideradas no estudo. A partir daí os dados poderiam ser ajustados para que um indicador não se sobrepusesse a outro, devido a magnitude dos seus valores.

Por fim, mas não menos importante, a escolha da configuração da rede neural e do modelo de compra e venda de ações completam a metodologia proposta, abrindo caminho para a fase de execução do projeto e consequente análise de resultados. O próximo capítulo traz de forma detalhada os resultados obtidos e as discussões acerca das predições realizadas. .

# Capítulo 4

## Resultados

O presente trabalho propõe estruturar um modelo baseado em redes neurais artificiais para prever preços de ativos do mercado de ações brasileiro. A predição tem como alvo as cotações de abertura dos 5 dias subsequentes à massa de dados, explorando para isso, indicadores baseados nas análises técnica e fundamentalista.

Para a realização do estudo, algumas etapas de desenvolvimento tiveram de ser seguidas: testes para definir a configuração de rede a ser utilizada, execução do modelo para todo o horizonte de predição, seleção do melhor modelo de compra e venda e por fim comparação com os benchmarks.

### 4.1 Requisitos do Sistema

Para construir o modelo baseado em redes neurais artificiais foi necessário um grande número de testes e simulações, para que, utilizando a estratégia de tentativa e erro, os melhores parâmetros fossem selecionados e o melhor ajuste fosse obtido.

Dessa forma foi necessário testar todos os arranjos descritos na sessão 3.2.1, para cada um dos dias-alvo e utilizando os dados referentes a cada uma das empresas. Para realização desse processo, utilizou-se do software Google Planilhas para organizar os dados extraídos do software Economática. Tais dados foram importados para o Google colab, ferramenta gratuita e online, onde todo o código foi escrito em linguagem Python.

A execução da coleta e organização dos dados foi realizada em um computador com sistema operacional Windows 10, processador Intel core i7 e 16 MB de memória RAM, já a execução do código responsável pela criação, treinamento, ajuste e testes da rede neural e cálculos de desempenho foi realizada em uma máquina virtual, disponibilizada pela ferramenta Google Colab, com 12 MB de memória RAM.

Rede	POCID	U-Theil	EMQ	R2
8	78,4%	0,90	0,73	96,8%
9	76,3%	0,93	0,68	97,9%
7	70,1%	1,02	0,98	97,5%
6	68,6%	1,91	1,77	91,3%
4	67,7%	1,89	1,89	90,7%
5	65,5%	2,33	2,60	87,1%
3	62,3%	3,32	2,56	79,7%
1	62,0%	3,12	3,88	68,1%
2	57,8%	3,90	6,67	64,5%

Tabela 4.1: Médias dos resultado das simulações de predição para “d+1”

## 4.2 Desenvolvimeto e Implementação

O primeiro passo do projeto foi definir qual a melhor configuração para a rede neural. Sendo assim uma simulação foi realizada com o objetivo de prever a cotação de abertura em “d+1” para o ativo BBDC4 utilizando os 9 modelos pré definidos na seção 3.2.1. Para cada modelo realizou-se 20 iterações uma vez que os pesos da rede foram iniciados de forma aleatória. A tabela 4.1 mostra as médias das métricas de avaliação de desempenho.

Analizando os resultados da tabela 4.1, que foram organizados em ordem decrescente do indicador POCID, pode concluir que a rede 8 obteve os melhores valores para as medidas de desempenho. Portanto definiu-se que os demais testes e simulações seriam realizados utilizando esta configuração de rede e os dados de entrada seriam selecionados com base nas maiores correlações.

A partir da escolha do melhor modelo e das variáveis de entrada foi necessário reproduzir o treinamento, testes e simulação para cada ativo utilizando todo o horizonte de predição proposto, “d+1”, “d+2”, “d+3”, “d+4” e “d+5”. Optou-se por, novamente, realizar um total de 20 iterações para cada alvo, ou seja, para cada ativo, foram realizadas 100 iterações, totalizando 1500 iterações para os 15 ativos. Os dados das predições e métricas de desempenho foram armazenados em planilhas para uma posterior análise.

A seguir ilustra-se esse procedimento exibindo as médias das medidas de desempenho para os testes feitos com o ativo BBDC4. Para tanto apresentam-se na tabela 4.2 os resultados alcançados para a fase de treinamento e a tabela 4.3 apresenta os resultados para a fase de simulação do modelo.

A partir da análise das tabelas 4.2 e 4.3 percebe-se que os valores das medidas de desempenho na fase de testes se assemelham aos da fase de simulação, confirmando assim a escolha da configuração da rede.



Projeção	POCID	U-Theil	EMQ	R2
d+1	78,4%	0,90	0,73	96,8%
d+2	57,7%	1,29	1,67	92,7%
d+3	53,3%	1,33	1,89	86,2%
d+4	56,1%	1,26	2,2	82,9%
d+5	53,7%	1,26	2,3	85,1%

Tabela 4.2: Médias dos resultado da fase de testes para todo o horizonte de predição

Projeção	POCID	U-Theil	EMQ	R2
d+1	79,1%	0,93	0,91	92,1%
d+2	55,2%	1,29	1,89	87,3%
d+3	51,3%	1,37	2,44	83,1%
d+4	55,3%	1,34	2,12	83,3%
d+5	50,7%	1,42	2,76	79,7%

Tabela 4.3: Médias dos resultado da fase de simulação para todo o horizonte de predição

## 4.3 Testes

O início da fase de testes do modelo de compra e venda se deu após o armazenamento de todas as 20 simulações realizadas para cada ativo. Como descrito na seção 3.3 o montante inicial seria de R\$10.000,00 e 2 estratégias de alocação do capital foram utilizadas. A seguir são apresentados os resultados para cada simulação.

### 4.3.1 Simulação 1

A primeira simulação foi realizada escolhendo aleatoriamente 1 entre as 20 iterações realizadas para cada ativo e a estratégia de investimento do capital definiu que R\$5.000,00 estariam disponíveis para investimento diário. Os resultados alcançados após 313 pregões são apresentados na tabela 4.4

A partir da análise da tabela 4.4 destacam-se os seguintes resultados: o maior lucro obtido foi de R\$251,77 com o ativo GGBR4 e o maior prejuízo foi de R\$148,10 com o ativo MULT3. Para 60% dos ativos o resultado foi positivo contra 40% dos ativos com resultado negativo. Resultado aparentemente positivo, porém o retorno total no período foi de apenas 3,8%

### 4.3.2 Simulação 2

Na simulação 2 a estratégia de compra e venda adotada definiu que todo o capital acumulado poderia ser utilizado para a compra de ações. Os resultados são apresentados na tabela 4.5

<b>Ativo</b>	<b>Retorno R\$</b>
BBDC4	-74,05
BRFS3	-3,19
BRKM5	-31,08
CSNA3	168,02
ELET3	0,38
EMBR3	-15,90
GGBR4	251,77
HAPV3	63,42
ITUB4	-27,66
KLBN4	45,87
MGLU3	57,89
MULT3	-148,10
OIBR4	42,48
PETR4	34,17
VALE3	13,32
<b>Total</b>	<b>377,34</b>

Tabela 4.4: Resultados da simulação 1

<b>Ativo</b>	<b>Retorno R\$</b>
BBDC4	-356,98
BRFS3	-79,57
BRKM5	-272,56
CSNA3	727,90
ELET3	-48,88
EMBR3	-203,32
GGBR4	986,02
HAPV3	301,18
ITUB4	-184,67
KLBN4	145,35
MGLU3	103,24
MULT3	-560,54
OIBR4	360,12
PETR4	39,02
VALE3	-0,76
<b>Total</b>	<b>955,54</b>

Tabela 4.5: Resultados da simulação 2

	Certo		Errado	
Ativo	Quant.	%	Quant.	%
BBDC4	65	47,1%	73	52,9%
BRFS3	62	48,1%	67	51,9%
BRKM5	53	45,3%	64	54,7%
CSNA3	88	50,9%	85	49,1%
ELET3	67	47,9%	73	52,1%
EMBR3	67	45,9%	79	54,1%
GGBR4	98	53,6%	85	46,4%
HAPV3	75	53,2%	66	46,8%
ITUB4	65	46,1%	76	53,9%
KLBN4	75	55,2%	61	44,8%
MGLU3	85	50,6%	83	49,4%
MULT3	51	44,4%	64	55,6%
OIBR4	46	51,7%	43	48,3%
PETR4	77	50,0%	77	50,0%
VALE3	94	52,2%	86	47,8%
<b>Total</b>	<b>1068</b>	<b>49,7%</b>	<b>1082</b>	<b>50,3%</b>

Tabela 4.6: Sinais de compra.

O resultado dessa vez foi melhor que o apresentado na simulação 1, porém ainda insatisfatório, uma vez que o retorno percentual alcançado foi de 9,6%, resultado abaixo do esperado. Destacam-se o melhor resultado, novamente para ativo GGBR4, incrementando o capital inicial em R\$986,02 e o pior resultado para MULT3 gerando um prejuízo de R\$560,54.

As tabelas 4.6 e 4.7 apresentam de forma detalhada todos os sinais de compra obtidos com a estratégia adotada para as simulações 1 e 2. Percebe-se que ao longo dos 313 pregões foram executados 2150 sinais de compra, dos quais 1068 resultaram em lucro e 1082 em prejuízo. Destaca-se também que o ativo KLBN4 obteve a maior taxa de acertos, 55,2%. O ativo com a menor taxa de acertos foi MULT3 com 44,3% dos sinais de compra resultando em lucro. A tabela 4.6 apresenta os acertos e erros para os sinais de compra e a tabela 4.7 os retornos obtidos para cada ativo.

### 4.3.3 Simulação 3

A fim de buscar a melhora dos resultados obtidos, para a simulação 3, optou-se por manter a estratégia de compra e venda da simulação 1, a qual estipulou o que o valor diário investido seria de R\$5.000,00, porém ao selecionar a rede a qual seria utilizada para realizar as previsões, utilizou-se a medida de desempenho POCID como base para a escolha, ou seja, seriam escolhidas as redes cuja fase de testes obteve o maior POCID entre as 20 iterações realizadas. Como explicado na seção 3.2.2 essa métrica

	Simulação 1			Simulação 2		
Ativo	Certo	Errado	Retorno	Certo	Errado	Retorno
BBDC4	404,10	-478,15	-74,05	1.670,33	-2.027,31	-356,98
BRFS3	525,67	-528,87	-3,19	2.057,56	-2.137,13	-79,57
BRKM5	509,73	-540,82	-31,08	1.954,48	-2.227,04	-272,56
CSNA3	804,99	-636,97	168,02	3.378,10	-2.650,20	727,90
ELET3	572,26	-571,88	0,38	2.272,17	-2.321,06	-48,88
EMBR3	609,63	-625,54	-15,90	2.423,30	-2.626,62	-203,32
GGBR4	818,59	-566,82	251,77	3.339,54	-2.353,52	986,02
HAPV3	517,61	-454,19	63,42	2.096,91	-1.795,73	301,18
ITUB4	389,15	-416,82	-27,66	1.593,70	-1.778,38	-184,67
KLBN4	678,51	-632,64	45,87	2.586,17	-2.440,82	145,35
MGLU3	709,81	-651,91	57,89	2.771,79	-2.668,55	103,24
MULT3	301,30	-449,40	-148,10	1.233,09	-1.793,63	-560,54
OIBR4	800,25	-757,76	42,48	3.472,31	-3.112,19	360,12
PETR4	541,72	-507,55	34,17	2.096,49	-2.057,48	39,02
VALE3	628,68	-615,36	13,32	2.500,60	-2.501,36	-0,76
<b>Total</b>	<b>8812,01</b>	<b>-8.434,67</b>	<b>377,34</b>	<b>35.446,57</b>	<b>-34.491,03</b>	<b>955,54</b>

Tabela 4.7: Retornos.

de desempenho indica os acertos do modelo quanto a direção da predição, ou seja, se o ativo irá se valorizar ou desvalorizar. A tabela 4.8 apresenta os resultados dessa simulação para cada ativo.

A utilização do POCID como balisador para escolha do modelo a ser utilizado fez com que o resultado melhorasse consideravelmente em relação a simulação 1, gerando um retorno de 15,3% contra 3,8% de forma que o resultado final gerou um lucro de R\$1.533,74. É possível notar que o número de ativos com resultado negativo reduziu-se a 3, ou seja apenas 20% do total de ativos. Ressalta-se também que os ativos com maior lucro e maior prejuízo se mantiveram em relação às simulações passadas, sendo eles GGBR4 e MULT3, respectivamente.

#### 4.3.4 Simulação 4

Por fim, com o objetivo de melhorar ainda mais o resultado optou-se por manter a estratégia de investimento da simulação 3, que permite que todo o capital seja investido, porém assim como na simulação 3, as redes utilizadas para as predições de cada ativo, seriam escolhidas com base no maior POCID dentre as 20 iterações feitas para cada dia alvo de cada ativo. A partir dessas definições, pode-se verificar os resultados na tabela 4.9.

Diante dos resultados apresentados anteriormente nota-se que a simulação 4 obteve o melhor resultado, com um retorno de 73% no período analisado o que equivale a um

Ativo	Retorno R\$
BBDC4	43,86
BRFS3	70,18
BRKM5	61,80
CSNA3	179,36
ELET3	191,09
EMBR3	76,34
GGBR4	392,81
HAPV3	128,88
ITUB4	-1,15
KLBN4	-82,74
MGLU3	275,13
MULT3	-133,20
OIBR4	178,03
PETR4	73,07
VALE3	80,29
<b>Total</b>	<b>1.533,74</b>

Tabela 4.8: Resultados da simulação 3

Ativo	Retorno R\$
BBDC4	124,46
BRFS3	325,10
BRKM5	25,62
CSNA3	1.084,85
ELET3	784,20
EMBR3	210,66
GGBR4	1.803,11
HAPV3	834,32
ITUB4	-43,71
KLBN4	-289,32
MGLU3	936,29
MULT3	-503,96
OIBR4	1.286,58
PETR4	229,84
VALE3	491,60
<b>Total</b>	<b>7.299,63</b>

Tabela 4.9: Resultados da simulação 4

Ativo	Certo		Errado	
	Quant.	%	Quant.	%
BBDC4	66	50,4%	65	49,6%
BRFS3	53	54,1%	45	45,9%
BRKM5	45	49,5%	46	50,5%
CSNA3	86	51,5%	81	48,5%
ELET3	78	56,5%	60	43,5%
EMBR3	67	50,8%	65	49,2%
GGBR4	98	58,7%	69	41,3%
HAPV3	69	53,5%	60	46,5%
ITUB4	56	48,3%	60	51,7%
KLBN4	74	57,4%	55	42,6%
MGLU3	89	52,7%	80	47,3%
MULT3	50	48,1%	54	51,9%
OIBR4	50	57,5%	37	42,5%
PETR4	79	53,0%	70	47,0%
VALE3	101	55,2%	82	44,8%
<b>Total</b>	<b>1061</b>	<b>53,3%</b>	<b>929</b>	<b>46,7%</b>

Tabela 4.10: Sinais de compra.

retorno de 57,8% a.a. As alterações feitas para as simulações 3 e 4 resultaram em uma série de mudanças quanto aos sinais de compra do modelo, começando pela redução de 7,5% no total de sinais. Outro ponto de extrema importância foi o incremento da acurácia dos sinais o que pode ser visto pela redução de 14,1% no sinais de compra errados.

As tabelas 4.10 e 4.11 apresentam de forma detalhada todos os sinais de compra obtidos com a estratégia adotada para as simulações 3 e 4. Percebe-se que ao longo dos 313 pregões foram executados 1990 sinais de compra, dos quais 1061 resultaram em lucro e 929 em prejuízo. Destaca-se também que o ativo GGBR4 obteve a maior taxa de acertos, 58,7%. O ativo com a menor taxa de acertos foi MULT3 com 48,1% dos sinais de compra resultando em lucro. A tabela 4.10 apresenta os acertos e erros para os sinais de compra e a tabela 4.11 os retornos obtidos para cada ativo.

## 4.4 Análise dos Resultados

A fim de facilitar o processo de avaliação das simulações e seus resultados, as principais informações a respeito das configurações do modelo e seus resultados foram compiladas e resumidas no quadro 4.12.

Após o processamento dos dados e realização das simulações é nítida a melhora dos resultados nas simulações 3 e 4, por consequência da utilização do POCID, não só

	Simulação 1			Simulação 2		
Ativo	Certo	Errado	Retorno	Certo	Errado	Retorno
BBDC4	410,50	-366,65	43,86	1.925,60	-1.801,14	124,46
BRFS3	416,79	-346,61	70,18	1.808,64	-1.483,55	325,10
BRKM5	505,76	-443,95	61,80	2.093,31	-2.067,69	25,62
CSNA3	836,22	-656,86	179,36	4.149,150	-3.064,30	1.084,85
ELET3	705,31	-515,22	191,09	3.039,27	-2.255,07	784,20
EMBR3	575,46	-499,12	76,34	2.685,73	-2.475,08	210,66
GGBR4	850,95	-458,14	392,81	3.919,53	-2.116,42	1.803,11
HAPV3	492,49	-363,61	128,88	2.470,42	-1.636,11	834,32
ITUB4	278,69	-279,84	-1,15	1.306,10	-1.349,81	-43,71
KLBN4	700,99	-783,72	-82,74	2.940,24	-3.229,56	-289,32
MGLU3	919,57	-644,44	275,13	3.949,92	-3.013,63	936,29
MULT3	315,62	-448,82	-133,20	1.423,06	-1.927,01	-503,96
OIBR4	911,21	-733,18	178,03	4.559,79	-3.273,21	1.286,58
PETR4	539,26	-466,19	73,07	2.360,52	-2.130,68	229,84
VALE3	718,13	-637,84	80,29	3.377,89	-2.886,29	491,60
<b>Total</b>	<b>9.176,93</b>	<b>-7.643,19</b>	<b>1.533,74</b>	<b>42.009,18</b>	<b>-34.709,55</b>	<b>7.299,63</b>

Tabela 4.11: Retornos.

Simul.	Configurações	Rent.	Rent. Anual
1	Limite de aplicação diário: R\$5.000,00. Seleção aleatória da rede	3,77%	2,98%
2	Limite de aplicação diário: Todo o montante. Seleção aleatória da rede	9,55%	7,57%
3	Limite de aplicação diário: R\$5.000,00. Seleção da rede com base no POCID	15,33%	12,15%
4	Limite de aplicação diário: Todo o montante. Seleção da rede com base no POCID	72,99%	57,83%

Tabela 4.12: Resumo das configurações e resultados das simulações

como métrica de avaliação de desempenho, mas também como balisador na escolha de qual rede utilizar. Essa decisão foi fundamental para o alcance das rentabilidades das simulações 3 e 4.

Percebe-se também que a configuração de compra e venda das simulações 1 e 3 são iguais, bem como as das simulações 2 e 4, tal mudança deve ser utilizada com cautela, pois apesar de aumentar os ganhos, também acentua as perdas, uma vez que o capital aplicado é maior, portanto caso o resultado geral seja negativo a decisão de disponibilizar todo o capital para investimento ocasionará em perdas ainda maiores.

## 4.5 Desempenho Comparativo

O objetivo da presente seção é realizar a comparação do modelo obtido com este estudo e as simulações realizadas com os benchmarks propostos e outros possíveis investimentos disponíveis no mercado financeiro brasileiro.

Para a realização das comparações escolheram-se os resultados da simulação 4, pois dentre as simulações feitas, essa foi a que apresentou os melhores resultados. A fim de manter exatamente as rentabilidades obtidas decidiu-se por, quando não houvesse nenhum sinal de compra em um determinado dia, nada seria feito, ou seja, nenhum rendimento seria acrescido ao montante, bem como foi feito durante a simulação.

Para dar inícios as comparações alguma premissa foram definidas. O inicial montante de todas as simulações foi de R\$10.000,00 e desconsiderou-se qualquer tipo de imposto sobre renda ou operação, bem como custos de corretagem e afins. Outra premissa importante do trabalho é que a execução da compra se daria sempre pela cotação de abertura do dia, algo que, na prática, pode não ser obtido com tamanha precisão.

A primeira comparação utiliza dados para o período de 08/10/2019 a 29/01/2021, desconsiderando finais de semana e feriados, formando portanto um espaço amostral de 313 pregões. Os ativos disponibilizados para investimentos são os mesmos utilizados neste estudo. Estes são apresentados na tabela 3.1.

Foram criadas 3 carteiras acionárias simuladas, a primeira delas teria metade do montante, ou seja, R\$5.000,00 disponíveis para investimento, valor o qual seria dividido igualmente para todos os ativos diariamente. A segunda carteira simulada poderia utilizar todo o montante - inclusive lucros, caso houvesse - para ser investido, dividindo tal valor, também igualmente entre os 15 ativos. A terceira carteira seguiu o formato de buy and hold, ou seja, o montante inicial foi dividido e investido igualmente entre todos os ativos e nenhuma venda foi realizada durante o período. Por fim a comparação também foi realizada com a rentabilidade do Ibovespa no mesmo período. Os retornos podem ser vistos na figura 4.1.



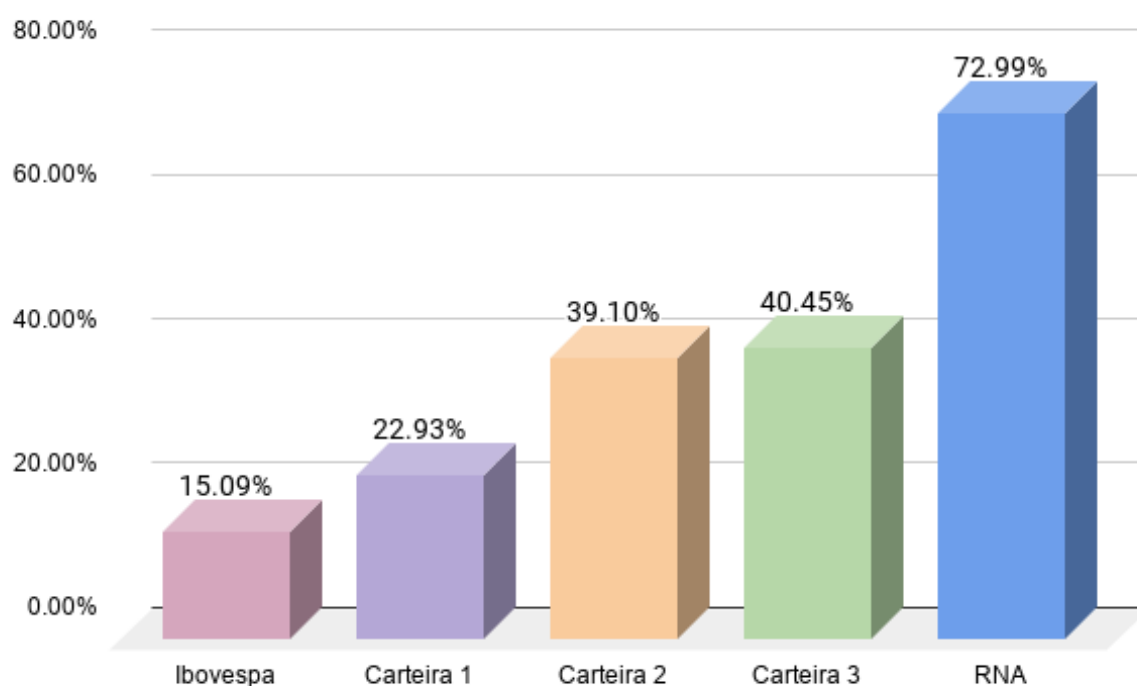


Figura 4.1: Rentabilidades dos banchmarks comparados à simulação 4. *Fonte:* Autor

Parâmetro	Carteira 3	RNA
Índice de Sharpe Diário	0,84	1,50
Máximo Drawdown	-52,51%	-33,77%
Melhor dia	13,69%	10,90%
Pior dia	-16,24%	-8,70%

Tabela 4.13: Resumo das configurações e resultados das simulações

Todas as rentabilidades foram positivas no período proposto, porém é notável a superioridade do modelo proposto, gerando um resultado quase 5 vezes ao do Ibovespa e quase 2 vezes à Carteira 3 que obteve o melhor resultado dentre os banchmarks. A figura 4.2 mostra detalhadamente os retornos do modelo e dos banchmarks durante o período analisado.

Nota-se que até a metade do período a rentabilidade acumulada do modelo esteve abaixo de pelo menos 1 dos banchmarks, porém a partir daí se destacou e assumiu a primeira colocação do ranking de rentabilidades.

A análise financeira sobre os banchmarks e o modelo podem se aprofundar um pouco mais para que, de forma numérica, se possa comprovar a vantagem de utilizar as predições realizadas pelo modelo somada a metodologia de alocação do capital como estratégia de investimento. Foram destacados 4 parâmetros os quais estão representados na tabela 4.13

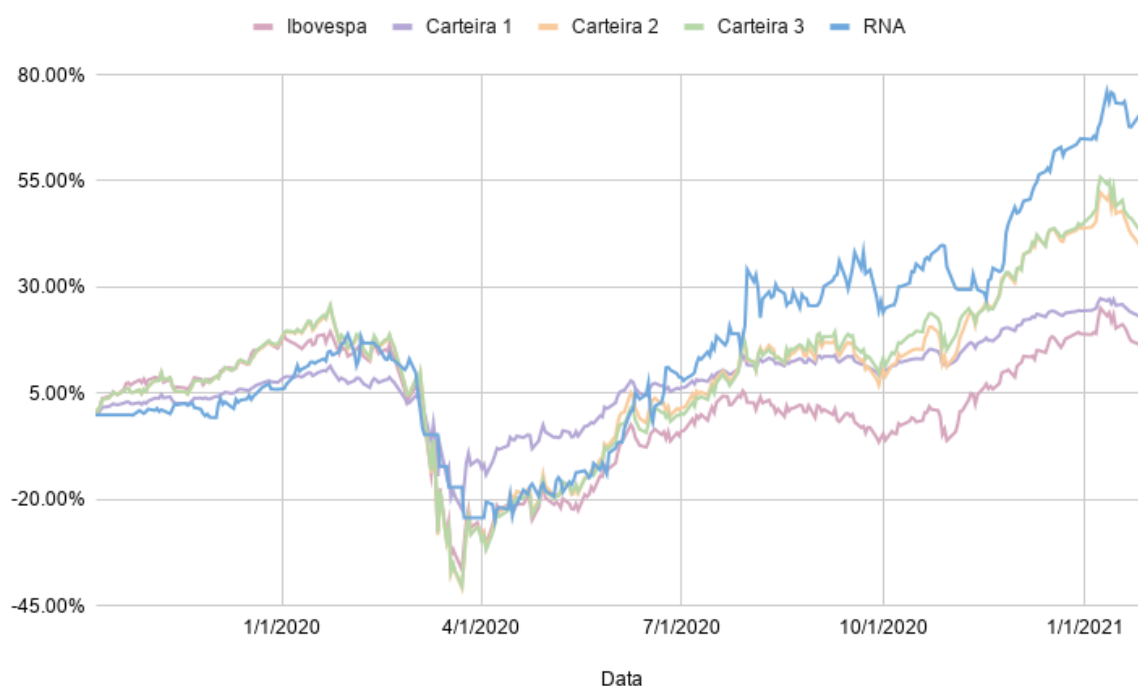


Figura 4.2: Rentabilidades dos banchmarks comparados à simulação 4 ao longo do tempo. *Fonte:* Autor

A carteira 3 foi escolhida para que a comparação fosse realizada, uma vez que, foi dela o melhor rendimento frente aos demais banchmarks. O índice de sharpe faz uma comparação entre o desempenho de um investimento em relação a outro com base no risco de se investir naquele portfólio, ou seja, mostra o quão bem o retorno de uma carteira de ativos compensa o investidor pelo risco assumido, portando quando maior melhor. Como mostrado na tabela 4.13 o índice de sharpe diário do modelo baseado em RNAs é quase o dobro em relação a carteira 3, logo, as decisões tomadas pelo modelo oferecem um melhor retorno para o mesmo risco ou, de forma equivalente, o mesmo retorno para risco menor.

A análise também permite a comparação entre os máximos drawdowns, indicador o qual mede a maior queda a partir da máxima histórica anterior. Na comparação entre o modelo e a carteira 3 é clara a queda mais acentuada quando a tomada de decisão não se baseia no modelo proposto, portanto pode-se inferir que além de gerar um retorno final superior, o modelo também tem um comportamento defensivo nas crises e em momentos de grandes quedas nas cotações.

É válido observar que a carteira 3 obteve, no seu melhor dia, um retorno 30% superior ao melhor dia do modelo, por outro lado, no pior dia a carteira 3 encerrou as posições no ativos com um prejuízo de quase 2 vezes o maior prejuízo obtido pelo

modelo.

Outra comparação que esse trabalho propõe-se a realizar é em relação aos fundos de investimentos disponíveis no mercado e que investem somente no mercado de ações brasileiro. Os fundos utilizados foram retirados dos sites InfoMoney e Valor Investe [13] [14] os quais foram classificados como os melhores fundos de investimentos no mercado brasileiro, em 2020 com base na sua rentabilidade. A figura 4.3 mostra a comparação entre os resultados dos fundos e do modelo.

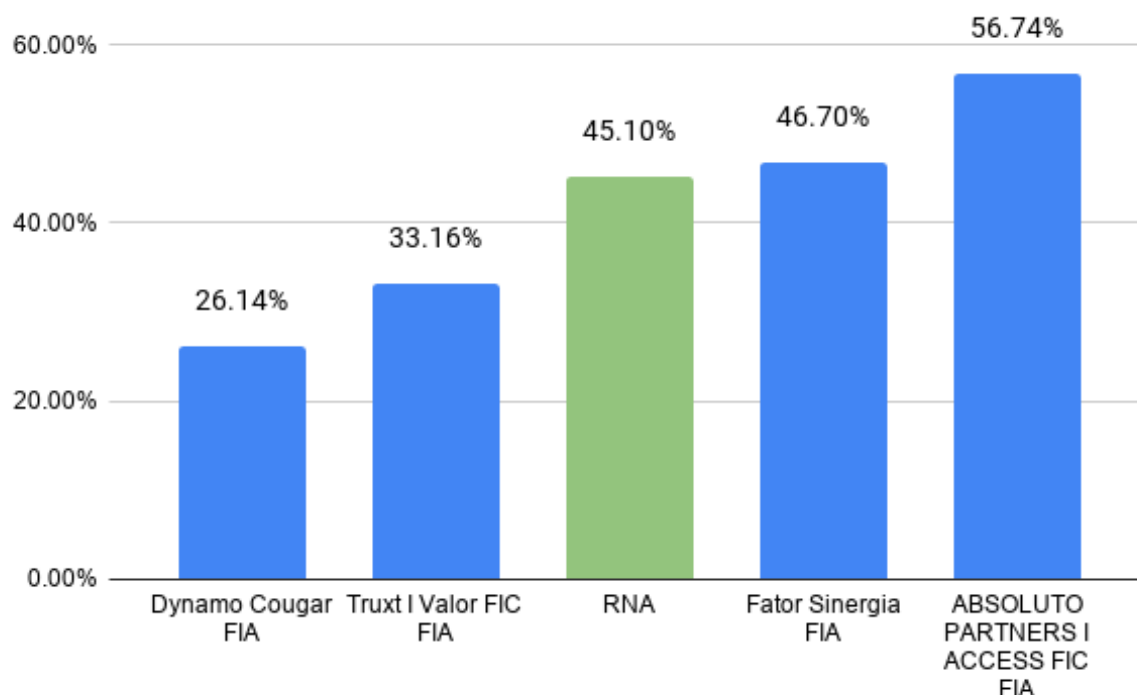


Figura 4.3: Rentabilidades dos fundos de ações e do modelo de redes neurais. *Fonte:* Autor

Note que a rentabilidade anual do modelo foi utilizada para a comparação e para que fosse suficientemente justa, foi aplicado um desconto mensal de 20% sobre os lucros, uma vez que esse é o imposto cobrado quando mais de R\$20.000,00 em vendas são realizadas no mês.

O rendimento obtido pelo modelo é surpreendente comparado aos fundos mais rentáveis de 2020, se assemelhando de forma notável ao fundo Fator Sinergia, o qual detém um patrimônio de aproximadamente 160 milhões de reais. Obteve também resultado superior aos fundos Dynamo Cougar e Truxt Valor FIC com rentabilidade de 33,16% e 26,14% e patrimônios de 1,02 bilhões e 7,68 bilhões respectivamente.



# Capítulo 5

## Conclusões

O presente trabalho buscou construir e apresentar um modelo baseado em redes neurais artificiais com o objetivo de auxiliar na tomada de decisão para a compra e venda de ativos no mercado financeiro brasileiro, além de construir uma estratégia adequada de investimento do capital.

Presumiu-se, com base nos estudos realizados, que as séries temporais ligadas aos ativos financeiros possuem memória de longo prazo, ou seja, seu movimento não é aleatório e podem ser identificados padrões de comportamento na variação dos preços. Dessa forma se torna possível estabelecer condições para que se utilizem dados passados para prever valores futuros dos preços dos ativos.

A seleção e pré-análise dos dados foi um fator crucial para o prosseguimento do estudo, uma vez que seriam essas as informações utilizadas para as previsões. A análise de correlação permitiu também que, dentre todos os indicadores selecionados, fosse utilizados somente aqueles que demonstraram ter um nível mínimo de relação com o alvo proposto.

O próximo passo do desenvolvimento foi estruturar um modelo de previsão baseado em redes neurais artificiais, o qual conseguisse manipular os dados, entender a dinâmica de variação dos preços e identificar as relações causais entre as variáveis de entrada e o alvo do modelo. Entretanto, foi nesse ponto onde começaram a surgir as dificuldades, já que, apesar do empolgante potencial apresentado pelas técnicas de finanças computacionais, em especial as RNAs, existe uma infinidade de configurações que podem ser assumidas pelo modelo além da extensa gama de possibilidades de combinações de dados de entrada do modelo.

Para contornar tais dificuldades, foram necessárias algumas escolhas as quais restringiram as infinitas possibilidades em um grupo menor, para que dessa forma, fosse possível explorar cada uma a fundo e escolher aquela que apresentasse os melhores resultados. Foram realizadas cerca de 2500 simulações, para que, além de obter os

resultados das predições de cada ativo, fosse possível realizar uma análise estatística e melhor fundamentar as decisões tomadas.

Todos os dados das simulações foram armazenados para uma posterior análise e assim realizar as melhores escolhas de configuração do modelo. Para mensurar o quão boas foram essas escolhas, estratégias de alocação do capital foram criadas para que, a partir dos sinais de compra gerados pelo modelo, o investimento pudesse ser realizado visando o maior retorno possível. Em suma foram realizadas quatro simulações, as quais combinaram a escolha das previsões com base na métrica de desempenho POCID, como as estratégias de alocação de capital.

A utilização do POCID como balisador da escolha do modelo a ser utilizado foi o grande diferencial para a melhoria dos resultados das duas últimas simulações, as quais apresentaram patamares de rentabilidade bastante significativos. Em comparação com a média de rentabilidade dos ativos e o investimento proporcional, seguindo as estratégias adotadas pelo modelo, a rentabilidade do modelo foi 80% superior as melhor banchmark, além de apresentar um retorno quase 5 vezes superior ao Ibovespa.

## 5.1 Considerações Finais

Assim sendo, acredita-se que o trabalho tenha contribuído para a área de finanças quantitativas, em especial para a utilização de redes neurais artificiais na predição de séries temporais relacionadas ao mercado financeiro.

Apresenta também, através das discussões e resultados apresentados, argumentos que suscitam o debate sobre a não aleatoriedade das séries financeiras, algo que contradiz um dos principais pilares da Teoria Financeira Moderna, que diz que tais séries temporais seguem um movimento browniano.

Por fim, promoveu-se um amplo estudo aplicado as redes neurais artificiais e principalmente no que diz respeito a modelagem dessa técnica. Assim, acredita-se que a presente pesquisa colabora para demonstrar o grande potencial das redes neurais quanto a sua utilização na área financeira, buscando estimular cada vez mais estudos e refinamento das técnicas utilizadas.

## 5.2 Propostas de Continuidade

Como proposta de continuidade e melhoria do trabalho, sugere-se que novas configurações de redes sejam testadas e que se utilize o modelo que melhor se adeque a cada ativo, e não somente um único modelo para todos os ativos.

Outra melhoria que influenciaria diretamente no resultado é a utilização de algoritmos de otimização no momento de alocação do capital. Dessa forma, quando mais de um sinal de compra fosse ativado para um determinado dia, o capital não fosse alocado igualitariamente, e sim de forma a priorizar um ativo com maior promessa de retorno.

Por fim, outra possibilidade é explorar novos horizontes de predição e diferentes estratégias de alocação de capital, além de incluir novas variáveis de entrada, as quais poderiam auxiliar na identificação de causalidades e consequentemente obter melhores predições.





# Referências Bibliográficas

- [1] MALKIEL, B. G. The valuation of closed-end investment-company shares. *The Journal of Finance*, 1977.
- [2] PAIVA, F. D. *Redes Neurais Para Decisões no Mercado de Ações Brasileiro*. [s.n.], 2014. Disponível em: <[http://repositorio.ufla.br/jspui/bitstream/1/2827/2/TESE\\_Redes](http://repositorio.ufla.br/jspui/bitstream/1/2827/2/TESE_Redes)>
- [3] BRAGA, A. de P. *Introdução à Engenharia de Dados Uma Perspectiva de Redes Neurais Artificiais e de Reconhecimento de Padrões*. [S.l.: s.n.], 2019.
- [4] FLOVIK, V. *How (not) to use Machine Learning for time series forecasting: Avoiding the pitfalls*. [s.n.], 2018. Disponível em: <<https://towardsdatascience.com/how-not-to-use-machine-learning-for-time-series-forecasting-avoiding-the-pitfalls-19f9d7adf424>>.
- [5] GOMES, F. *A Bolsa de Valores brasileira como fonte de informações financeiras*. [s.n.], 1997. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/633/422>>.
- [6] MOREIRA, F. *B3 (B3SA3): volume financeiro em 2020 é de R\$ 6,45 tri e se aproxima do PIB*. Disponível em: <<https://www.euqueroinvestir.com/b3-b3sa3-volume-financeiro-aproxima-pib/>>.
- [7] FLECK, L. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, 2016. Disponível em: <<https://periodicos.utfpr.edu.br/recit/article/view/4330/Leandro>>.
- [8] OLAH, C. Understanding lstm networks. *Colah's Blog*, 2015. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>.
- [9] BINHURAIB, T. The essence of rnns. *Towards Data Science*, 2020. Disponível em: <<https://towardsdatascience.com/the-essence-of-rnns-44dfb4107a47>>.
- [10] NIMESH, S. Understanding lstm and its quick implementation in keras for sentiment analysis. *Towards Data Science*, 2018. Disponível em: <<https://towardsdatascience.com/understanding-lstm-and-its-quick-implementation-in-keras-for-sentiment-analysis-af410fd85b47>>.

- [11] GUIMARÃES, P. R. B. Análise de correlação e medidas de associação. *Universidade Federal do Paraná. Disponível em:* <<https://docs.ufpr.br/~jomarc/correlacao.pdf>>. Acesso em, v. 9, 2017.
- [12] REDUÇÃO de Variáveis de Entrada de Redes Neurais Artificiais a Partir de Dados de Análise de Componentes Principais na Modelagem de Oxiô.
- [13] BOMBANA, L. *Os melhores (e os piores) fundos de ações e multimercados em dezembro e em 2020*. Disponível em: <<https://www.infomoney.com.br/onde-investir/os-melhores-e-os-piores-fundos-de-acoes-e-multimercados-em-dezembro-e-em-2020/>>.
- [14] BERTÃO, N. *Ranking: Conheça os fundos de ações e 'long biased' que mais renderam em 2020*. Disponível em: <<https://valorinveste.globo.com/produtos/fundos/acoes/noticia/2021/01/08/fundos-de-investimento-em-acoes-e-long-biased-chegam-a-render-mais-de-80percent.ghml>>.