

CENTRO UNIVERSITÁRIO CARIOCA

FERNANDO PASSOS MELLO  
RAMON RAMALHO CAMPOS

O USO DA INTELIGÊNCIA ARTIFICIAL NA ANÁLISE FUNDAMENTALISTA PARA  
TOMADA DE DECISÃO NO MERCADO DE AÇÕES

Rio de Janeiro  
2023

FERNANDO PASSOS MELLO  
RAMON RAMALHO CAMPOS

O USO DA INTELIGÊNCIA ARTIFICIAL NA ANÁLISE FUNDAMENTALISTA PARA  
TOMADA DE DECISÃO NO MERCADO DE AÇÕES

Trabalho de Conclusão de Curso apresentado ao Centro  
Universitário Carioca, como requisito a obtenção do  
grau de Bacharel em Ciência da Computação.

Orientadora: Prof<sup>a</sup>. Daisy Cristine Albuquerque da Silva

Rio de Janeiro  
2023

FERNANDO PASSOS MELLO  
RAMON RAMALHO CAMPOS

O USO DA INTELIGÊNCIA ARTIFICIAL NA ANÁLISE FUNDAMENTALISTA PARA  
TOMADA DE DECISÃO NO MERCADO DE AÇÕES

Trabalho de Conclusão de Curso apresentado ao Centro  
Universitário Carioca, como requisito a obtenção do  
grau de Bacharel em Ciência da Computação.

Cidade, \_\_ de \_\_\_\_\_ de \_\_\_\_

BANCA EXAMINADORA

---

Prof. Dr. ....

Universidade .....

---

Prof. Dr. ....

Universidade .....

## RESUMO

O presente trabalho tem como foco o uso da Inteligência Artificial (IA) na análise fundamentalista para tomada de decisão no mercado de ações. O principal objetivo demonstrar e avaliar um modelo de IA capaz de auxiliar na tomada de decisão no mercado de ações por meio da análise fundamentalista. Para tal, foi realizada uma revisão bibliográfica abrangente acerca do mercado de ações, análise fundamentalista e IA, além da criação de um modelo de IA específico para este objetivo. A justificativa para o uso da IA na análise fundamentalista reside na importância dessa abordagem no mercado financeiro, visando otimizar a tomada de decisão e impactar empresas, investidores e o mercado de ações em geral. Isso se dá através de decisões mais informadas, rápidas e estratégicas, minimização de erros humanos, processamento de grandes volumes de dados, avaliação de riscos, e identificação de padrões e tendências. O estudo é caracterizado como descritivo, com resultados avaliados qualitativa e quantitativamente, baseados na coleta de fontes secundárias e dados de empresas listadas na bolsa de valores brasileira. A análise dessas informações levou à conclusão de que, embora o modelo de IA criado não tenha alcançado rentabilidade positiva no período testado, os resultados fornecem informações valiosas sobre os desafios da implementação de estratégias de investimento baseadas em IA no mercado de ações, além de destacar a necessidade de considerar abordagens mais abrangentes e variadas.

**Palavras-chave:** Inteligência artificial; análise fundamentalista; mercado de ações; tomada de decisão.

## **ABSTRACT**

This work focuses on the use of Artificial Intelligence (AI) in fundamental analysis for decision-making in the stock market. The main objective is to demonstrate and evaluate an Artificial Intelligence (AI) model capable of assisting in decision-making in the stock market through fundamental analysis. To this end, a comprehensive literature review was carried out on the stock market, fundamental analysis and AI, in addition to the creation of a specific AI model for this objective. The justification for using AI in fundamental analysis lies in the importance of this approach in the financial market, improving decision making and impacting companies, investors and the stock market in general. This happens through more informed, faster and strategic decisions, minimization of human errors, processing of large volumes of data, risk assessment, and identification of patterns and trends. The study is characterized as descriptive, with results evaluated qualitatively and quantitatively, based on the collection of secondary sources and data from specific companies on the Brazilian stock exchange. Analysis of this information led to the conclusion that although the created AI model did not succeed in positive profitability in the tested period, the results provide excellent insights into the challenges of implementing AI-based investment strategies in the stock market, as well as highlight the need to consider more comprehensive and varied approaches.

**Keywords:** Artificial intelligence; fundamental analysis; stock market; decision-marking.

## SUMÁRIO

|         |   |    |
|---------|---|----|
| 1       | INTRODUÇÃO .....  | 6  |
| 2       | REFERENCIAL TEÓRICO .....                               | 8  |
| 2.1     | MERCADO DE AÇÕES .....                                  | 8  |
| 2.1.1   | ANÁLISE FUNDAMENTALISTA .....                           | 9  |
| 2.2     | INTELIGÊNCIA ARTIFICIAL .....                           | 11 |
| 2.2.1   | APRENDIZADO DE MÁQUINA .....                            | 13 |
| 2.3     | INTELIGÊNCIA ARTIFICIAL NA ANÁLISE FUNDAMENTALISTA .... | 14 |
| 3       | DEMONSTRAÇÃO DO USO DE IA NO MERCADO DE AÇÕES .....     | 16 |
| 3.1     | LÓGICA DO PROJETO .....                                 | 16 |
| 3.2     | SELEÇÃO DAS EMPRESAS .....                              | 16 |
| 3.3     | EXTRAÇÃO DOS DADOS .....                                | 16 |
| 3.3.1   | FUNDAMENTOS DAS EMPRESAS .....                          | 16 |
| 3.3.2   | COTAÇÕES E IBOVESPA .....                               | 17 |
| 3.4     | TRANSFORMAÇÃO DOS DADOS .....                           | 17 |
| 3.4.1   | FUNDAMENTOS .....                                       | 18 |
| 3.4.2   | COTAÇÕES .....  | 18 |
| 3.4.3   | FUNDAMENTOS E COTAÇÕES .....                            | 19 |
| 3.4.4   | IBOVESPA .....  | 20 |
| 3.4.5   | APLICAÇÃO DA ANÁLISE FUNDAMENTALISTA .....              | 20 |
| 3.4.5.1 | INDICADORES .....                                       | 21 |
| 3.4.5.2 | REMOÇÃO DE VALORES NULOS E DE COLUNAS .....             | 21 |
| 3.4.5.3 | UNINDO OS DADOS DAS EMPRESAS .....                      | 21 |
| 3.4.6   | ANÁLISE EXPLORATÓRIA .....                              | 22 |
| 3.4.6.1 | RESPOSTAS POR VIÉS .....                                | 22 |
| 3.4.6.2 | ANÁLISE DE CORRELAÇÃO .....                             | 24 |
| 3.5     | SELEÇÃO DE CARACTERÍSTICAS .....                        | 26 |
| 3.6     | CRIAÇÃO DA INTELIGÊNCIA ARTIFICIAL .....                | 27 |
| 3.6.1   | STANDARD SCALER .....                                   | 27 |
| 3.6.2   | DADOS DE TREINO E DADOS DE TESTE .....                  | 27 |
| 3.6.3   | DUMMY CLASSIFIER .....                                  | 28 |
| 3.6.4   | MODELOS DE TESTE .....                                  | 29 |
| 3.6.5   | AJUSTE DE PARÂMETROS .....                              | 32 |
| 3.7     | AVALIAÇÃO .....   | 34 |
| 3.7.1   | AVALIAÇÃO DA INTELIGÊNCIA ARTIFICIAL .....              | 34 |
| 4       | RESULTADOS .....  | 36 |
| 4.1     | AVALIAÇÃO COMPARATIVA .....                             | 36 |
| 5       | CONCLUSÃO .....   | 38 |
|         | REFERÊNCIAS .....                                       | 39 |

## 1 INTRODUÇÃO

O presente projeto de pesquisa tem como foco principal abordar o uso da Inteligência Artificial (IA) no mercado de investimentos, especificamente no mercado de ações, como auxílio para tomada de decisão com base em uma análise fundamentalista. Portanto, é importante aprofundarmos a compreensão sobre como a IA está transformando o mercado de investimentos, destacando sua influência nas decisões financeiras e na avaliação de empresas negociadas na bolsa de valores.

Segundo Fabrício Gonçalves (2023), CEO da *Box Asset*, no site Capital Aberto, em um artigo sobre como a IA poderá transformar os investimentos, afirma que a IA está revolucionando o mundo, especialmente no mercado de ações, e que é inevitável reconhecer que ela está impactando cada vez mais as avaliações de empresas negociadas na bolsa de valores. Sendo assim, no contexto do mercado financeiro e de investimentos, o uso da IA tem ganhado destaque na otimização dos processos decisórios, especialmente no mercado de ações. O advento e a evolução das tecnologias digitais trouxeram consigo uma revolução na forma como as decisões financeiras são tomadas, possibilitando a análise de grandes volumes de dados e a automação de processos antes realizados manualmente.

Em vista dessa revolução proporcionada pela IA no mercado financeiro e suas contribuições na otimização dos processos decisórios, é relevante e conveniente investigar mais detalhadamente o papel da IA na análise fundamentalista, que é essencial na tomada de decisão para investimentos de longo prazo no mercado de ações.

Discutir sobre o uso da IA na análise fundamentalista para tomada de decisão no mercado de ações justifica-se por ser um assunto de crescente importância no contexto do mercado financeiro e na busca pela otimização dos processos de decisão. De acordo com a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), as empresas financeiras têm adotado cada vez mais a IA para se beneficiarem da enorme quantidade de conjuntos de dados disponíveis. Assim, é possível notar que o uso da IA no processo decisório pode impactar direta ou indiretamente empresas, investidores e o mercado financeiro como um todo, através de uma tomada de decisão mais informada, mais rápida e mais concentrada na estratégia. Além disso, impacta na redução de erros humanos, no processamento de grandes volumes de dados, na avaliação de riscos e na identificação de padrões e tendências. Para tanto, é necessário entendermos como funciona a IA e a análise

fundamentalista no contexto do mercado de investimentos e demonstrar como se dá essa relação por meio uma aplicação.

Sendo assim, o presente trabalho, em conformidade com o problema de pesquisa, estabeleceu como objetivo geral demonstrar e avaliar o uso de um modelo de IA, com base na análise fundamentalista, capaz de auxiliar na tomada de decisão no mercado de ações. Para alcançar o objetivo geral, serão abordados tópicos sobre o mercado de ações, análise fundamentalista, conceitos de IA, aprendizado de máquina e seus algoritmos, além de uma abordagem específica por meio da criação de um modelo de IA capaz de construir uma carteira de ações utilizando a análise fundamentalista.

Portanto, o presente estudo consiste em uma pesquisa aplicada de caráter descritivo, que visa demonstrar e avaliar o uso da IA na análise fundamentalista para tomada de decisão no mercado de ações. Nesse sentido, os resultados serão apresentados de forma qualitativa e quantitativa, a partir da coleta de informações de fontes secundárias, como livros, outras produções acadêmicas, sites, blogs, entre outras fontes. Além disso, também serão utilizados dados de indicadores financeiros e fundamentalistas de empresas listadas na B3.

Para alcançar o objetivo central, este trabalho encontra-se organizado em 5 capítulos, sendo esta introdução o primeiro deles. No Capítulo 2, tratamos do referencial teórico. No Capítulo 3, demonstramos e avaliamos o uso da IA no mercado de ações capaz de construir uma carteira de ações com base em uma análise fundamentalista das empresas. Os resultados serão abordados no Capítulo 4. E, por fim, no Capítulo 5, são apresentadas as considerações finais.



## 2 REFERENCIAL TEÓRICO

### 2.1 MERCADO DE AÇÕES

Ao longo do tempo, o mercado de ações tem sido um tópico extensivamente estudado por vários autores, resultando em uma vasta quantidade de literatura que engloba diversas definições. No entanto, existem certos conceitos que se destacam significativamente dentro dessa área. Segundo Tiago Reis, especialista em investimentos e fundador da Suno Research, no portal Suno, em artigo sobre Bolsa de valores, Mercado de ações: como funciona? Entenda tudo sobre, afirma que:

O mercado de ações é o ambiente de negociações onde investidores, através das bolsas de valores, realizam operações de compra e de venda de suas participações acionárias de empresas que possuem capital aberto (as famosas S/A, ou Sociedades Anônimas). Isto além de títulos de dívidas dessas empresas e outros ativos relacionados às mesmas.

Outra abordagem seria conforme o quarto capítulo do guia POR DENTRO DA BM&FBOVESPA, Guia prático de uma das maiores bolsas de valores do mundo, em sua 1ª edição, publicado pelo INSTITUTO EDUCACIONAL BM&FBOVESPA, que afirma:

O mercado de renda variável é composto por títulos e valores mobiliários cujo valor varia ao longo do tempo apenas em função das condições de oferta e demanda e não em função de variáveis ou indicadores pré-estabelecidos. Considerando essa ampla definição, compõem o mercado de renda variável as ações emitidas pelas companhias de capital aberto, os índices de ações, os fundos de ações, certificados emitidos sobre ações e fundos de investimento imobiliários, entre outros.

As ações são a menor parcela do capital aberto de uma empresa, sendo estes títulos de renda variável. De acordo com Pinheiro (2008), existem critérios adotados para diferenciar ações, no qual destaca-se a espécie da ação, sendo definida pelas vantagens e direitos que detém. Desta forma, a ação pode ser preferencial ou ordinária, tendo como principal diferença entre as duas o direito a voto em assembleias e prioridade de recebimento de dividendos. Uma ação é um valor mobiliário, expressamente previsto no inciso I, do artigo 2º, da Lei nº 6.385/1976.

As abordagens analisadas sobre o mercado de ações oferecem diferentes perspectivas. Tiago Reis apresenta uma visão geral, caracterizando o mercado de ações como um ambiente de negociações para compra e venda de participações acionárias de empresas com capital

aberto e outros ativos relacionados. O Instituto Educacional BM&FBOVESPA enfatiza a volatilidade do mercado de renda variável, composto por títulos cujo valor flutua devido à oferta e demanda, incluindo uma variedade de títulos além das ações. Por último, Pinheiro se aprofunda nas características específicas das ações, destacando critérios de diferenciação, como a distinção entre ações preferenciais e ordinárias, e seus direitos associados.

Considerando as análises e definições previamente estabelecidas, pode-se inferir que o mercado de ações, independentemente da terminologia atribuída, configura-se como um ambiente público altamente organizado destinado à negociação de ativos financeiros emitidos por empresas de capital aberto. Todas as abordagens são complementares e contribuem para uma compreensão mais completa do mercado de ações.

Aprofundando o entendimento sobre o mercado de ações, é fundamental conhecer as estratégias de investimento disponíveis para maximizar os rendimentos. Uma das principais abordagens é a análise fundamentalista. Essa técnica é muito útil para investidores que desejam identificar oportunidades de investimento de médio e longo prazo, proporcionando uma análise rigorosa que leva em consideração diversos fatores financeiros e econômicos da empresa.

### **2.1.1 ANÁLISE FUNDAMENTALISTA**

A estratégia conhecida como análise fundamentalista, cujo objetivo é determinar o valor intrínseco das ações e, conseqüentemente, avaliar seu preço justo, foi desenvolvida ao longo do século XX, pelo economista Benjamin Graham. Essa abordagem se apoia nas características financeiras da empresa, sua perspectiva de crescimento, perfil de risco e outros indicadores relevantes. Com base nessa análise, os investidores fundamentais tomam decisões de compra ou venda de ativos, sendo mais apropriada para operações de médio e longo prazo.

PINHEIRO (2008), afirma que:

É a análise que utiliza os fundamentos econômico-financeiros da empresa para determinar o seu valor justo. Assim, a análise fundamentalista avalia e tenta encontrar alternativas de investimento no mercado acionário através da avaliação das conjunturas macro e microeconômicas, além do panorama setorial da empresa e das demonstrações financeiras e balanços.

Ainda segundo Pinheiro (2008), a Avaliação de uma ação à luz da análise fundamentalista pode ser feita observando 3 pontos:

1. Confiabilidade do investidor em razão do tamanho da influência governamental;
2. Demonstrativos financeiros, como também informações importantes, que possibilitem uma projeção da evolução da empresa emissora da ação; e
3. A economia do mercado como um todo, estimativas futuras que tenham influência sobre o preço de todas as ações negociadas.

Todavia, cabe ressaltar que há outros que também são de suma importância, sendo estes:

- lucro esperado de exercício;
- investimentos realizados e futuros; e
- fontes de financiamentos.

De acordo com Braga (2009), a análise fundamentalista se destaca por sua abordagem na análise dos fundamentos das empresas, empregando uma ampla variedade de informações para auxiliar os investidores na tomada de decisões.

Para Cavalcante et al (2005), a análise fundamentalista é uma estratégia de avaliação que visa determinar o valor adequado de uma ação com base nas perspectivas futuras de desempenho da empresa. Isso ocorre porque o preço das ações no mercado de ações reflete as emoções e expectativas médias dos investidores. O valor da ação negociada no mercado é, em essência, um reflexo do sentimento do mercado, enquanto o valor adequado representa uma avaliação individual da verdadeira condição da empresa.

Segundo Noronha (1995 apud PIAZZA, 2008, p.23):

É a ciência que busca, por meio do estudo de registros multiformes, associados a formulações matemático-estatísticas, incidente sobre preços, volumes e contratos em aberto do passado e do corrente dos diferentes ativos financeiros, proporcionar, pela análise de padrões que se repetem condições para que possamos projetar o futuro caminho dos preços, dentro de uma lógica de maiores probabilidades.

Entretanto, Braga (2009), destaca que, não basta somente o acesso à informação, é necessário o uso correto para que o processo de avaliação considere o maior número de informações disponíveis, para que efetivamente possa apoiar o investidor na determinação do “preço” da empresa.

De acordo tais abordagens, a análise fundamentalista é uma estratégia de avaliação de ações baseada nas características financeiras da empresa, perspectivas de crescimento, perfil de risco e outros indicadores relevantes. Ela utiliza os fundamentos econômico-financeiros da empresa para determinar seu valor justo, considerando conjunturas macro e microeconômicas, panorama setorial, demonstrações financeiras e balanços. A avaliação de uma ação por meio dessa abordagem pode incluir a confiabilidade do investidor, informações financeiras e projeções de evolução da empresa, e a economia do mercado como um todo. A análise fundamentalista se destaca por sua abordagem nos fundamentos das empresas e emprega uma ampla variedade de informações para auxiliar os investidores na tomada de decisões. Contudo, é essencial o uso correto dessas informações para que o processo de avaliação seja eficaz na determinação do preço da empresa.

A análise fundamentalista tem sido uma ferramenta valiosa para os investidores na avaliação de ações e na tomada de decisões de investimento. No entanto, à medida que a tecnologia avança, novas abordagens e métodos vêm ganhando espaço no mercado financeiro. Um dos avanços mais significativos nos últimos anos tem sido a aplicação da IA nos investimentos, auxiliando, automatizando e otimizando muitos aspectos nesse processo.

## 2.2 INTELIGÊNCIA ARTIFICIAL

A emergência do conceito de IA na história remonta aos anos 1950, no período pós-Segunda Guerra Mundial. Nesse contexto, pesquisadores do campo da ciência da computação começaram a questionar a viabilidade de desenvolver sistemas de computação capazes de realizar atividades cognitivas, emulando a capacidade de 'pensar' dos seres humanos. Inicialmente, a IA foi definida como o campo de pesquisa que trata da 'automação de tarefas intelectuais normalmente desempenhadas por seres humanos', conforme afirmado por Chollet (2017).

Essa definição inicial serviu como ponto de partida para compreender por que a IA abrange áreas como *Machine Learning* (Aprendizado de Máquina) e *Deep Learning* (Aprendizado Profundo), embora existam outros paradigmas que não se encaixem no modelo de aprendizado. Um exemplo clássico são os primeiros programas de xadrez, que eram codificados com regras do jogo e simulavam posições sem envolver nenhum processo de aprendizado.

De acordo com Coelho (2020), a IA é um campo da Ciência da Computação que se concentra na automação do comportamento inteligente, abrangendo uma ampla gama de pesquisas. Um desses subcampos é o Aprendizado de Máquina, que emprega algoritmos e metodologias para aprender com dados históricos.

A IA é capaz de simular não apenas o comportamento humano, mas também sua capacidade cognitiva. Isso é alcançado por meio de algoritmos que podem analisar grandes volumes de dados em um curto período e aprender continuamente, permitindo assistência na tomada de decisões. A análise de dados, aliada à capacidade de aprendizado dos algoritmos, possibilita a solução rápida de problemas sem intervenção humana, como destacado por Lobo (2018).

Coelho (2020) argumenta que o Aprendizado de Máquina é a capacidade de um computador aprender sem ter sido programado explicitamente para uma tarefa, dependendo apenas de exemplos passados. Isso demonstra que, devido à capacidade de aprendizado dos algoritmos, é plenamente possível desenvolver sistemas capazes de executar tarefas humanas de forma automatizada.

Segundo Russell; Norvig (2004), a IA pode ser conceituada como o campo que reside na interseção entre ciência e arte, com seu foco central sendo a resolução de problemas de maneira análoga à inteligência humana.

A integração da análise fundamentalista com a IA representa um avanço notável na esfera da tomada de decisões de investimento no mercado de ações. A análise fundamentalista tradicional desempenha um papel fundamental, fornecendo uma base sólida para avaliar a saúde financeira das empresas, considerando indicadores financeiros essenciais, tais como lucratividade, balanços patrimoniais e projeções de crescimento. A IA, por sua vez, aprimora essa análise ao viabilizar o processamento eficaz de vastos volumes de dados em tempo real, identificando correlações complexas que frequentemente escapam à percepção humana. Dessa forma, a sinergia entre essas abordagens possibilita uma avaliação mais abrangente e precisa das empresas, enriquecendo substancialmente a qualidade das decisões de investimento e ampliando as oportunidades de êxito no dinâmico contexto do mercado de capitais.

Após examinarmos o impacto significativo da IA na análise de informações financeiras e no processo de tomada de decisões no contexto do mercado de ações, nosso enfoque se volta agora para uma subdivisão crucial desse campo: o Aprendizado de Máquina.

### 2.2.1 APRENDIZADO DE MÁQUINA

O Aprendizado de Máquina é um subcampo da IA que se caracteriza por programas capazes de adquirir conhecimento com base em experiências previamente registradas. De acordo com Weiss e Kulikowski (1991), uma definição amplamente aceita de AM é a seguinte: "É um sistema computacional que toma decisões com base em exemplos e experiências anteriores".

Este domínio da IA pode ser subdividido em dois principais paradigmas:

1. **Aprendizado Supervisionado:** Este paradigma visa identificar e modelar a relação entre as entradas (inputs) e as saídas (outputs) fornecidas, classificando-as e rotulando-as em categorias pré-definidas, como imagens, documentos, entre outros tipos de dados.
2. **Aprendizado Não Supervisionado:** O Aprendizado Não Supervisionado tem como objetivo identificar padrões e agrupar elementos sem rótulos claros. Entre suas diversas aplicações, destaca-se a classificação textual; e
3. **Aprendizado por Reforço:** Concentra no desenvolvimento de sistemas de aprendizado autônomo, nos quais um agente interage com um ambiente, toma ações e recebe feedback na forma de recompensas.

Conforme observado por Faceli et al. (2011), a Aprendizagem de Máquina é uma área da IA dedicada ao desenvolvimento de programas computacionais capazes de executar tarefas e aprender com base em sua própria experiência. Isso envolve o desenvolvimento de algoritmos que podem aprender de forma autônoma, usando informações e dados de experiências passadas. Este campo de pesquisa multidisciplinar abrange diversas disciplinas, incluindo Ciência da Computação, Probabilidade e Estatística, Teoria da Complexidade, Teoria da Informação, entre outras áreas de conhecimento.

Em contrapartida, Burkov (2019) enfatiza que nenhuma máquina possui a verdadeira capacidade de aprendizado. Em termos simplificados, o que uma máquina comumente realiza é a busca por uma fórmula matemática que, quando alimentada com um conjunto de dados (denominado dados de treinamento), seja capaz de gerar um tipo específico de saída.

Segundo Bengio Yoshua; Courville (2016), a aplicação prática da IA, bem como do Aprendizado de Máquina, é desafiada por uma série de fatores que influenciam a qualidade e

relevância dos dados observados, o que, por sua vez, impacta as avaliações sobre o que é ou não relevante no contexto da IA e do AM.

Em síntese, a Aprendizagem de Máquina representa um subcampo crucial da IA, permitindo que programas computacionais aprendam com base em experiências prévias e dados. Compreender essas nuances é essencial para explorar o potencial da IA e do AM, bem como para enfrentar os obstáculos que surgem ao aplicá-los em cenários do mundo real.

## **2.3 INTELIGÊNCIA ARTIFICIAL NA ANÁLISE FUNDAMENTALISTA**

A incorporação da IA no mercado de ações, em paralelo à análise fundamentalista, tem se tornado uma prática comum. A utilização da IA se torna de grande relevância devido à capacidade significativa de processamento em conjunto com um grande volume de dados. Além do poder de processamento, a habilidade de avaliação de riscos também contribui para a geração de informações relevantes durante o processo de investimento.

No contexto da análise fundamentalista, a IA demonstra competência ao examinar dados financeiros cruciais, como balanços, fluxos de caixa e retornos sobre ativos. Por outro lado, a abordagem envolve o uso de séries temporais financeiras, aplicando métodos matemáticos e modelos baseados em regras para embasar decisões informadas. É viável treinar um sistema de IA para identificar empresas com desempenho sólido.

De acordo com Tiago Reis (2023), CEO do portal Suno, os benefícios proporcionados pela IA ao mercado financeiro estão relacionados ao aprimoramento da precisão nas previsões, eficiência operacional, personalização de serviços financeiros e à capacidade de analisar grandes volumes de dados. Esses dados são obtidos por meio da análise fundamentalista das empresas.

Além disso, a IA desempenha um papel fundamental na automação de tarefas rotineiras e repetitivas, como a execução de ordens de compra e venda, monitoramento de carteiras e cálculos de risco. A automação de processos representa um benefício significativo que a IA pode proporcionar às empresas do mercado financeiro, aos profissionais da área e aos investidores individuais.

Dessa forma, diversas ferramentas de IA podem ser empregadas, ampliando exponencialmente as capacidades de um operador, seja processando um grande volume de dados que anteriormente demandaria muito tempo, realizando transações de compra e venda de ações ou construindo uma carteira de investimentos.

A integração da IA na análise fundamentalista estabelece uma base sólida para a exploração das potencialidades da IA no mercado financeiro. Ao enfatizar a capacidade da IA de examinar dados financeiros cruciais e aprimorar a avaliação de riscos, a análise fundamentalista se torna um ponto de partida estratégico. Essa transição naturalmente nos conduz à próxima fase: a demonstração prática do uso da IA no mercado de ações.



### **3 DEMONSTRAÇÃO DO USO DE IA NO MERCADO DE AÇÕES**

#### **3.1 LÓGICA DO PROJETO**

Este trabalho consiste em criar uma IA capaz de analisar os dados de indicadores financeiros de várias empresas e tomar uma decisão de compra ou venda das ações. Ou seja, essa IA será capaz de montar uma carteira de ações inteligente, que consiga analisar os fundamentos das empresas e, com base nessa análise, decidir quais empresas comprar ou vender. Além disso, essa tomada de decisão ocorrerá dentro de um determinado período, trimestre por trimestre. Ao final, avaliamos a rentabilidade da carteira criada pela IA e comparamos com o índice Ibovespa.

#### **3.2 SELEÇÃO DAS EMPRESAS**

Para selecionarmos as empresas que fariam parte deste projeto, recorreremos aos índices da B3, que reúne importantes indicadores de desempenho das ações das empresas listadas na bolsa de valores. Entre estes índices, foram selecionadas todas as empresas listadas na carteira teórica do IBrX 100, válida para o quadrimestre de setembro a dezembro de 2023. Este índice tem o objetivo de ser o indicador do desempenho médio das cotações dos 100 ativos de maior negociabilidade e representatividade do mercado de ações.

#### **3.3 EXTRAÇÃO DOS DADOS**

##### **3.3.1 FUNDAMENTOS DAS EMPRESAS**

Após a seleção das empresas, foi necessário buscar os dados dos seus fundamentos, o balanço patrimonial e o demonstrativo de resultados de cada uma dessas empresas. A fonte de dados escolhida foi o sistema online do site Fundamentus, que oferece informações financeiras e fundamentalistas de empresas listadas na B3. Este sistema permite o download (feito de forma manual) de um arquivo de determinada empresa no formato .zip, contendo uma pasta de trabalho do Excel com duas planilhas, sendo uma com os dados históricos trimestrais de balanço patrimonial e a outra com os dados históricos trimestrais do demonstrativo de resultados (DRE). Esses arquivos foram extraídos para uma pasta específica e renomeados para que fossem identificados nos passos posteriores do projeto.

A planilha de balanço patrimonial possui os dados sobre o patrimônio, a liquidez, a rentabilidade e a estrutura de capital da empresa. Cada coluna representa dados específicos sobre o ativo total, o passivo total e o patrimônio líquido. Já a planilha de demonstrativo de resultados contém dados sobre as receitas, os custos, as despesas e o lucro da empresa. Cada coluna também representa um trimestre e cada linha representa os dados específicos de receitas, custos, despesas e lucro.

### 3.3.2 COTAÇÕES E IBOVESPA

Para obtermos o histórico de cotações das ações, foi utilizado como fonte de dados o Yahoo Finance, um site que oferece uma série de recursos para investidores e profissionais do mercado financeiro, inclusive as cotações de ações.

Por meio de uma biblioteca de código aberto, a *yfinance*, que facilita a extração de dados financeiros do Yahoo Finance, foi possível obter o histórico dos dados de cotações de cada uma das empresas que foram selecionadas anteriormente.

Os dados obtidos nesse processo de extração resultam em um *Data Frame*, uma estrutura de dados bidimensional do Pandas, que é uma biblioteca de manipulação e análise de dados muito popular do Python. Cada linha desse *Data Frame* representa uma data diferente, e as colunas contêm informações sobre os preços de abertura (*Open*), o preço mais alto atingido (*High*), o preço mais baixo atingido (*Low*), o preço de fechamento (*Close*), o preço de fechamento ajustado (*Adj Close*) e o número de ações negociadas (*Volume*).

Além dos dados de cotações das empresas, também foi extraído o histórico do Ibovespa, um índice que representa a performance média do mercado de ações brasileiro e que nos permitirá, ao final do projeto, fazer uma avaliação comparativa com a rentabilidade da carteira de ações criada pela IA.

## 3.4 TRANSFORMAÇÃO DOS DADOS

A etapa de transformação dos dados é uma parte crucial para o processo de preparação dos dados para análise. Essa etapa envolve a manipulação e o ajuste dos dados brutos para que eles estejam em um formato adequado para análise.

### 3.4.1 FUNDAMENTOS

Para cada empresa, foram lidas duas planilhas distintas, sendo uma contendo os dados de balanço patrimonial e outra contendo os dados de demonstrativo de resultados. Cada uma dessas planilhas foi lida como um *Data Frame*, sendo o código da ação como índice. Em seguida, os *Data Frames* referentes a cada uma das planilhas foram concatenados. Por fim, esses *Data Frames* concatenados foram armazenados em um dicionário, com a chave sendo o código do ativo correspondente.

### 3.4.2 COTAÇÕES

Um dicionário também foi criado com o propósito de armazenar as cotações de cada empresa. Para cada empresa, foi adicionada uma chave no dicionário, correspondendo ao código da ação dessa empresa. O valor associado a essa chave é um subconjunto do *Data Frame* de cotações criado anteriormente, contendo exclusivamente as linhas que estão associadas à empresa em questão.

Neste dicionário de cotações criado, algumas datas trimestrais estavam ausentes, devido ao fato de essas datas corresponderem a finais de semana ou feriados. Para resolver essa questão, foi necessário incluir essas datas faltantes no dicionário. A abordagem utilizada foi refletir o valor da cotação da data anterior para as datas ausentes. Para tanto, foi criado um índice de datas trimestrais que cobre o período entre 01/01/2013 e 31/12/2022. Em seguida, para cada empresa presente no dicionário, as datas trimestrais que estavam ausentes foram preenchidas no *Data Frame* de cotações associado à respectiva empresa.

Após o processo de inclusão das datas trimestrais ausentes, foi observado que ainda poderiam existir linhas vazias no *Data Frame* de cotações. Essa situação poderia ocorrer caso algumas empresas não possuíssem cotação disponível no Yahoo Finance ou caso não tivessem uma base de dados suficientemente grande para a análise. Para verificar essa situação, foi feita uma verificação da presença de valores nulos no *Data Frame* de cotações associado a cada empresa. As empresas que apresentassem valores nulos foram removidas do dicionário, visando garantir a consistência dos dados. Após essa remoção, restaram 71 empresas das 100 iniciais para prosseguir com o projeto.

### 3.4.3 FUNDAMENTOS E COTAÇÕES

Primeiramente, foi realizada a mesclagem dos dicionários de dados de cada empresa, ou seja, os dados de fundamentos e os dados de cotação de cada empresa foram mesclados. Nesse processo, o *Data Frame* de fundamentos da empresa passou por uma transposição, trocando linhas por colunas, e o índice da tabela foi convertido para o tipo *datetime*. Em seguida, foi criado um *Data Frame* contendo apenas a coluna “Adj Close” do *Data Frame* de cotações da empresa, que era a coluna de interesse. Posteriormente, os dois *Data Frames*, o de fundamentos e o de cotações, foram mesclados utilizando seus índices como chave para a mesclagem. Por fim, o dicionário de fundamentos foi atualizado com a versão modificada do novo *Data Frame*.

Foi observado que algumas empresas possuíam colunas diferentes em seus *Data Frames*, o que poderia impactar a análise. Sendo assim, foi definido um vetor com os nomes das colunas do *Data Frame* da empresa “ABEV” para servir como referência para a comparação com os outros *Data Frames* no dicionário. Com base nisso, as empresas que tinham colunas diferentes das colunas de referência foram removidas do dicionário. Após essa remoção, restaram 66 empresas das 100 iniciais para continuar o projeto. Além disso, foi identificado que algumas colunas tinham nomes repetidos no *Data Frame*, sendo o mesmo nome de coluna para os dados de balanço patrimonial e para os dados de demonstrativo de resultados. Assim, foi necessário alterar os nomes dessas colunas para que fossem diferenciadas e tivessem nomes únicos. Para isso, foi feita uma verificação nas colunas de referência para identificar quais estavam duplicadas. As colunas duplicadas foram renomeadas, adicionando um sufixo “(1)” em seu nome, e os nomes das colunas de todos os *Data Frames* no dicionário de fundamentos das empresas foram atualizados.

Também foi identificado que algumas colunas possuíam valores vazios, o que poderia comprometer a análise. Dessa forma, as colunas que tinham valores vazios ou grande parte deles foram removidas. Para isso, foram calculados o número total de valores nulos para cada coluna em todos os *Data Frames* do dicionário de fundamentos, bem como o número total de linhas nesses *Data Frames*. A partir dessa análise, foi possível observar que algumas colunas estavam com valores nulos em todas as linhas, de um total de 2617 linhas. Além disso, foi identificada uma coluna com uma quantidade de 74 valores nulos. Mesmo sendo um número baixo em relação ao total de linhas, para evitar qualquer tipo de influência negativa na análise,

as colunas que tinham mais de 70 valores nulos foram removidas de todos os *Data Frames*, e os valores nulos remanescentes foram preenchidos com o último valor válido.

#### 3.4.4 IBOVESPA

Assim como ocorreu com o Data Frame das cotações das ações, o Data Frame do Ibovespa também apresentava datas trimestrais ausentes. Diante disso, foi realizada a inclusão das datas trimestrais que estavam faltando no Data Frame do Ibovespa.

Posteriormente, foi realizada a mesclagem dos dados com o índice Ibovespa. Esse processo consistiu na inclusão, em cada Data Frame do dicionário de fundamentos, de uma nova coluna contendo os dados do índice Ibovespa correspondentes às datas presentes no respectivo Data Frame.

#### 3.4.5 APLICAÇÃO DA ANÁLISE FUNDAMENTALISTA

Em nossa análise fundamentalista, não poderíamos avaliar os indicadores das empresas apenas como números brutos, uma vez que cada empresa tem indicadores e números distintos, que podem ter significados diferentes para cada uma delas. Como solução, optamos por transformar esses números em percentuais de variação.

Para classificar e determinar, com auxílio de IA, se deveríamos comprar, não comprar ou vender uma ação, comparamos os relatórios de cada trimestre. A partir de um relatório trimestral, analisamos as mudanças nos números em relação ao trimestre passado e observamos a variação de cotação em relação ao trimestre seguinte. Isso nos permitiu determinar se houve uma indicação de compra, não compra ou venda de uma ação, observando se as variações dos números de um trimestre para o seguinte afetam os preços futuros.

Para garantir que não estaríamos criando regras aleatórias, comparamos o desempenho das ações com o índice Ibovespa, um indicador que representa o desempenho médio das principais ações na bolsa de valores brasileira. Dessa forma, conseguimos avaliar o desempenho de uma ação em relação ao mercado geral.

As regras estabelecidas foram as seguintes:

- Se a ação subiu mais do que o Ibovespa (ou caiu menos), deveríamos comprar (2).

- Se a ação subiu menos do que o Ibovespa até o Ibovespa - 2% (ou caiu mais do que o Ibovespa até Ibovespa - 2%), deveríamos não comprar (1).
- Se a ação subiu menos do que o Ibovespa - 2% (ou caiu mais do que o Ibovespa - 2%), deveríamos vender (0).

#### 3.4.5.1 INDICADORES

Primeiramente, foram estabelecidas condições para calcular os retornos das colunas do Data Frame, bem como uma lista de valores para calcular os retornos com base nessas condições.

Posteriormente, os retornos da cotação (“*Adj Close*”) e do Ibovespa (IBOV) foram calculados como variação percentual em relação ao valor seguinte.

Em seguida, foi definida outra lista de condições para calcular a coluna viés. Essa coluna de viés foi calculada com base nos valores selecionados a partir dessas condições. Por fim, o dicionário de fundamentos foi atualizado com a versão modificada.

#### 3.4.5.2 REMOÇÃO DE VALORES NULOS E DE COLUNAS

Após a transformação dos dados, foi feito um levantamento do número total de linhas em todos os *Data Frames*, bem como do número de valores nulos em cada coluna. O resultado desse processo foi que todas as colunas, com exceção da coluna de viés, apresentavam diversas linhas vazias. Essa situação ocorreu porque, antes de transformarmos os indicadores em percentuais, havia muitas linhas com valores zerados. Diante disso, de um total de 2617 linhas, foi decidido então remover as colunas que tinham mais de um terço de valores nulos em todos os *Data Frames* no dicionário de fundamentos. Os valores nulos remanescentes foram preenchidos com valores zerados. Além disso, como as colunas “*Adj Close*”, “IBOV” e “Resultado” não seriam mais relevantes para a nossa análise, também as removemos do dicionário.

#### 3.4.5.3 UNINDO OS DADOS DAS EMPRESAS

Antes de unirmos os *Data Frames* de todas as empresas, realizamos uma cópia do dicionário dos dados de fundamentos das empresas. Essa medida foi tomada para garantir que

as alterações feitas durante o processamento e a concatenação dos *Data Frames* não afetassem o dicionário original. Esse procedimento é uma prática importante para assegurar a integridade dos dados originais.

Em seguida, removemos a primeira e a última linha de cada *Data Frame*, visto que a primeira linha não poderia ser utilizada para avaliar os fundamentos, já que eles são baseados em dados passados, e a última linha não poderia ser utilizada para avaliar a cotação, já que as cotações são baseadas em dados futuros.

Também redefinimos o índice dos *Data Frames*, ou seja, os índices de datas foram excluídos e substituídos por números, uma vez que essas datas não seriam mais relevantes para nossa análise.

Por fim, concatenamos todos os *Data Frames* em um único *Data Frame*, que passou a ser nossa base de dados principal para análise.

### 3.4.6 ANÁLISE EXPLORATÓRIA

Com o intuito de entendermos as características e propriedades dos dados, foi feita uma análise exploratória da nossa base de dados. A análise exploratória nos ajuda a ter uma melhor compreensão dos dados, a identificar padrões e tendências, a definir a estratégia de pré-processamento, a selecionar as características e validar suposições. Ou seja, é uma etapa fundamental no processo de análise de dados, pois fornece uma base sólida para tomada de decisões e construção de modelos preditivos.

#### 3.4.6.1 RESPOSTAS POR VIÉS

Na nossa análise exploratória, foi realizada uma avaliação da quantidade de respostas por viés. Para visualizar a distribuição dos valores da coluna “Vies” do *Data Frame* de nossa base de dados, utilizamos a biblioteca *plotly.express*. Contamos a frequência de cada valor único da coluna “Vies” e exibimos a distribuição normalizada no formato de porcentagem.

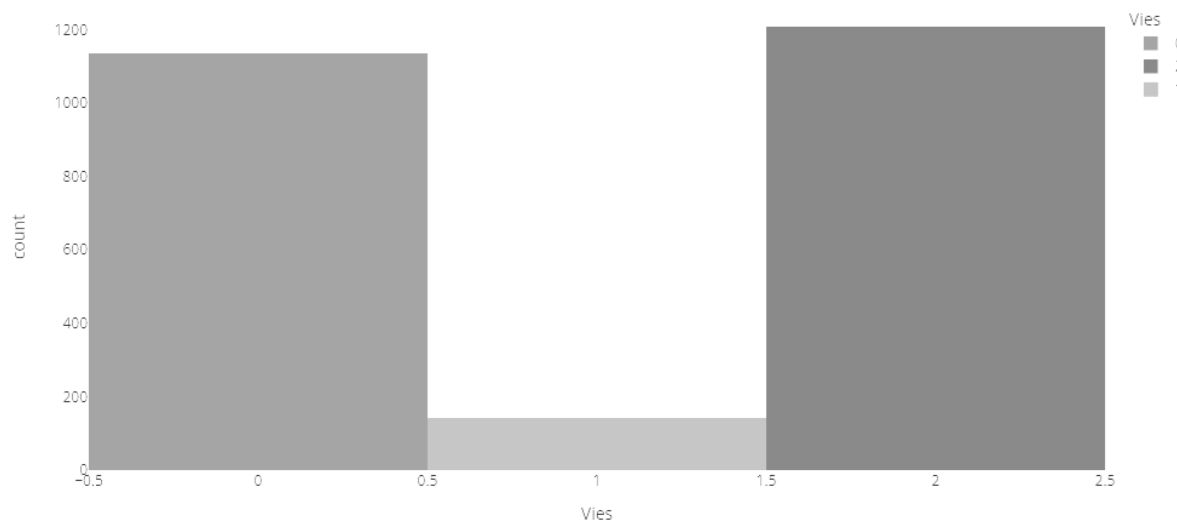


Gráfico 1: Histograma de Respostas por Viés, sendo o viés 0 para vender, 1 para não comprar e 2 para comprar.

O resultado dessa análise foi a exibição da distribuição normalizada dos valores da coluna “Viés” no formato de porcentagem, bem como um histograma que ilustra a distribuição dos valores de “Viés”. Foi possível observar que a frequência do viés com valor 2 (comprar) foi de 48,6% (1207), a frequência do viés com valor 1 (não comprar) foi de 5,7% (142) e a frequência do viés com valor 0 (vender) foi de 45,7% (1136).

Diante disso, foi possível notar que a frequência do viés com valor 1 (não comprar) foi significativamente menor e desproporcional em comparação às outras duas categorias. Por esse motivo, decidimos ajustar os valores que sinalizam “não comprar” (1) para “vender” (0), eliminando assim essa categoria da análise. A partir desse ponto, passamos a concentrar nossa atenção nos vieses iguais a 0 (vender) e 2 (comprar).

Para implementar essa mudança, localizamos todas as ocorrências da coluna “Viés” de valor 1 (não comprar) e atribuímos a elas o valor 0 (vender). Em seguida, exibimos novamente a distribuição normalizada dos valores da coluna “Viés” no formato de porcentagem e um histograma para visualizarmos a alteração.



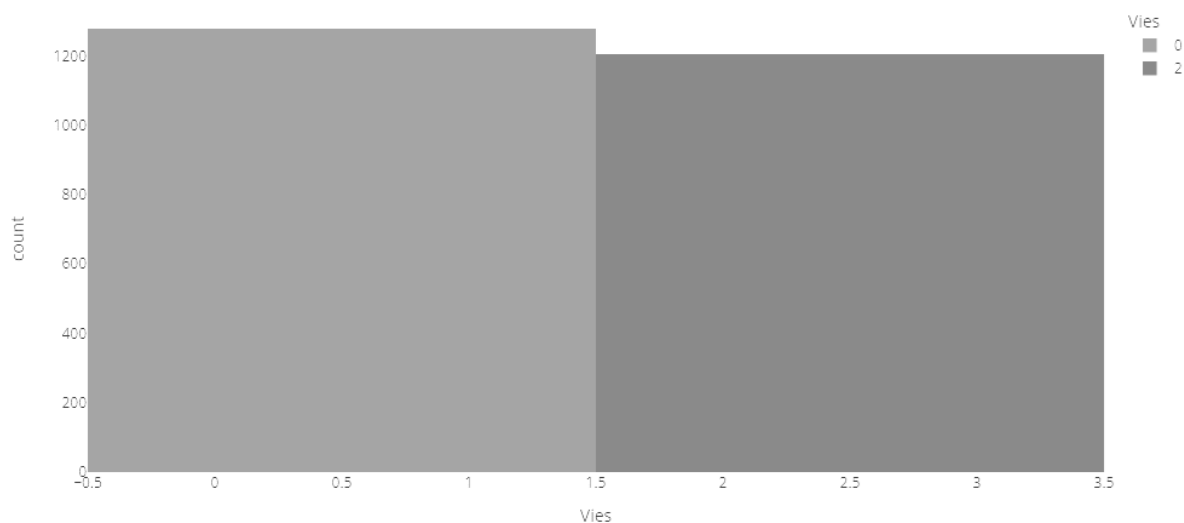


Gráfico 2: Histograma de Respostas por Viés, sendo o viés 0 ou 1 para vender e 2 para comprar.

Após essa mudança, a frequência do viés com valor 0 (vender) passou a ser de 51,4% (1278) e a frequência do viés com valor 2 (comprar) se manteve em 48,6% (1207).

### 3.4.6.2 ANÁLISE DE CORRELAÇÃO

Realizamos uma análise de correlação com o objetivo de entender a relação linear entre duas variáveis quantitativas, ou seja, como a mudança em uma variável está relacionada à mudança na outra. Essa análise nos auxilia a identificar as variáveis que possuem uma relação forte, o que é crucial para selecionarmos as características relevantes e construir o nosso modelo.

Calculamos a matriz de correlação entre todas as combinações de pares de colunas do Data Frame da nossa base de dados. Em seguida, criamos e exibimos um gráfico de *heatmap* dessa matriz de correlação, o qual ilustra as correlações existentes entre as colunas do Data Frame da nossa base de dados.

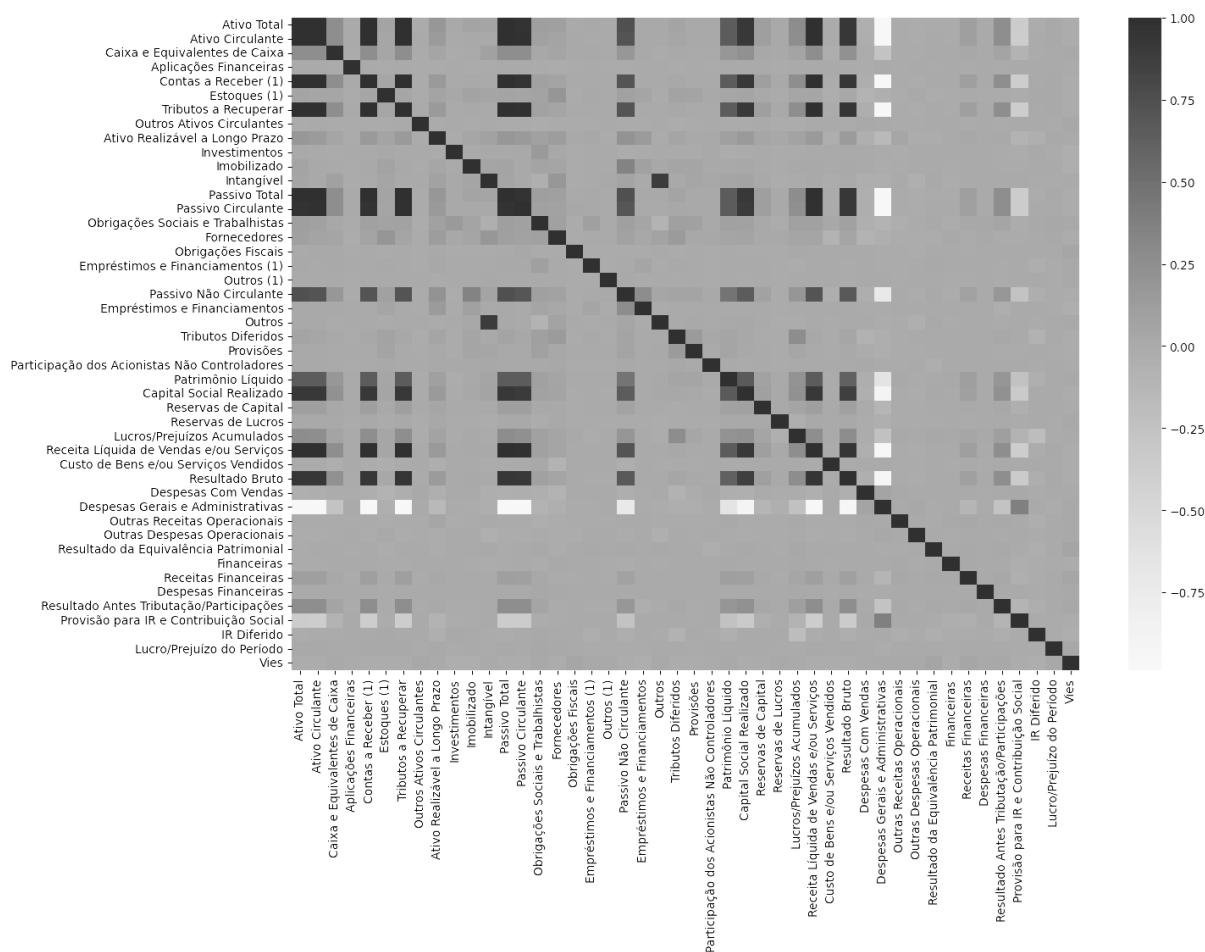


Gráfico 3: Gráfico de calor da correlação entre as variáveis.

A partir do gráfico de *heatmap*, foi possível observar uma correlação muito forte entre diversas variáveis. Esse fenômeno pode ser problemático, pois dificulta a identificação das variáveis que são realmente importantes para a modelagem, uma vez que o efeito de uma variável pode ser mascarado pelo efeito de outra variável correlacionada. Por isso, decidimos remover as variáveis altamente correlacionadas.

Para identificar quais variáveis remover, procuramos por pares de colunas na nossa base de dados que tinham uma correlação absoluta maior que 0,8. A partir disso, observamos que várias correlações possuíam uma correlação muito forte com a coluna “Ativo Total”. Dessa forma, optamos por remover essas variáveis e manter somente a variável “Ativo Total”. Essa abordagem foi adotada porque, ao remover essas variáveis, podemos reduzir o risco de que elas também apresentem correlações muito fortes entre si.

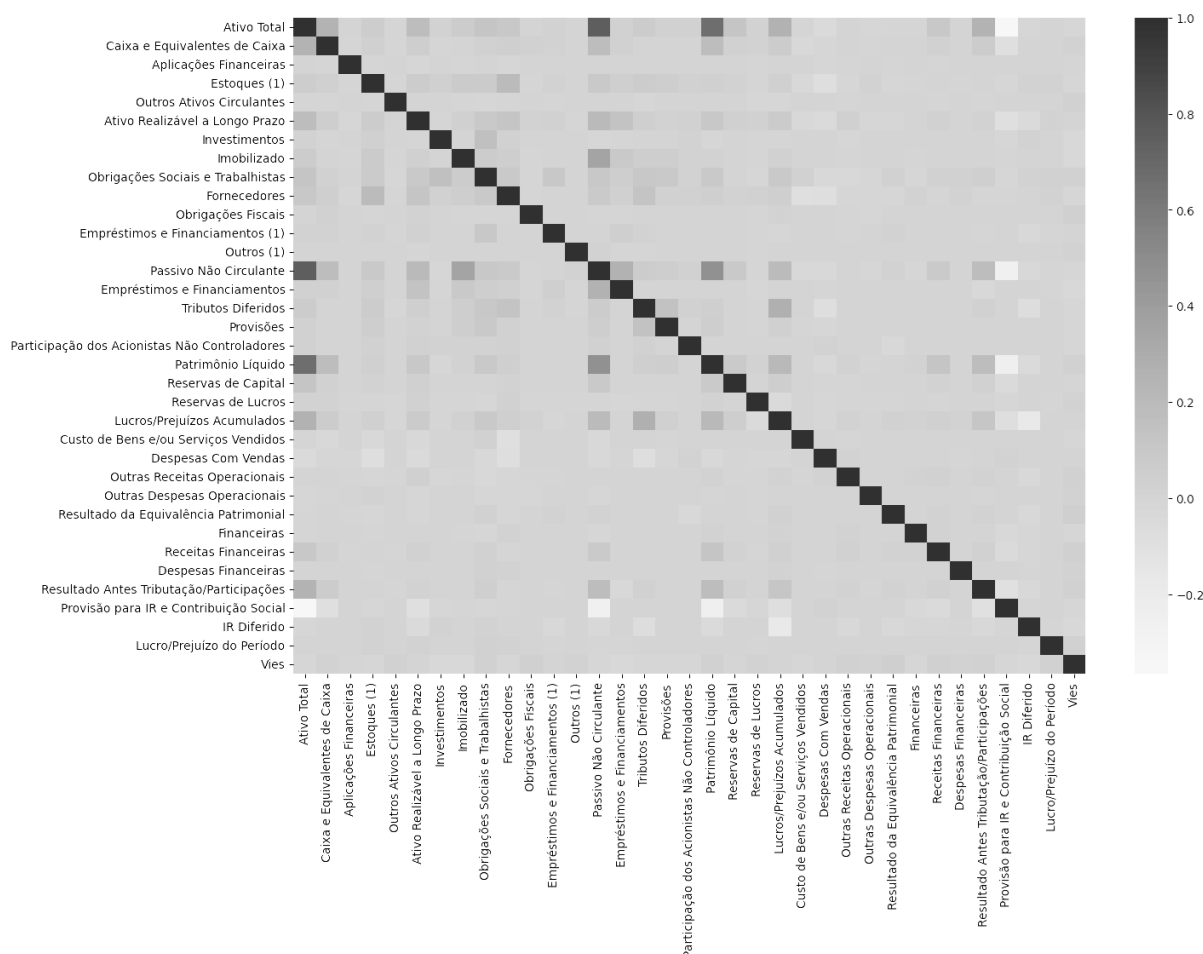


Gráfico 4: Gráfico de calor da correlação entre as variáveis após as alterações.

Ao recalcular a matriz de correlação do Data Frame e exibir um novo gráfico de *heatmap* dessa matriz, foi possível observar uma diminuição significativa na quantidade de variáveis com correlações muito fortes.

### 3.5 SELEÇÃO DE CARACTERÍSTICAS

A etapa de *Feature Selection*, ou Seleção de Características, tem como propósito identificar as variáveis mais relevantes para o modelo que será construído. Esse processo pode otimizar a performance do modelo, reduzir a complexidade e diminuir o tempo de treinamento. No entanto, ter um número elevado de variáveis pode ser prejudicial, pois, à medida que o número de variáveis aumenta, a quantidade de dados necessária cresce exponencialmente. Como não dispomos de uma quantidade de dados tão grande quanto a

quantidade de variáveis, o ideal seria reduzir o número de variáveis sem comprometer a eficiência.

Para selecionar as melhores características e construir o modelo com base nessa seleção reduzida, treinamos uma árvore de decisão com os dados da nossa base de dados. Considerando que nossa base de dados não é muito grande, é perfeitamente viável treinar uma IA simples, pois o processo não seria demorado.

A partir desse treinamento, identificamos as 10 características mais importantes para o nosso modelo: Lucros/Prejuízos Acumulados, Provisão para IR e Contribuição Social, Obrigações Fiscais, Lucro/Prejuízo do Período, Outros Ativos Circulantes, IR Diferido, Fornecedores, Obrigações Sociais e Trabalhistas, Resultado Antes Tributação/Participações e Receitas Financeiras.

### 3.6 CRIAÇÃO DA INTELIGÊNCIA ARTIFICIAL

#### 3.6.1 STANDARD SCALER

StandardScaler é uma técnica de pré-processamento de dados que tem como objetivo padronizar e normalizar as características em um conjunto de dados. Foi exatamente isso que fizemos com a nossa base de dados. Para aplicar essa técnica, definimos uma função que padroniza as características utilizando o StandardScaler. Em seguida, aplicamos essa função, obtendo como resultado um Data Frame contendo as 10 características mais importantes, todas padronizadas, além da coluna "Vies".

#### 3.6.2 DADOS DE TREINO E DADOS DE TESTE

Dividimos o nosso Data Frame em conjuntos de treinamento e teste, com o objetivo de treinar e avaliar o nosso modelo de *machine learning*. Como resultado, obtivemos quatro conjuntos de dados: “x\_treino” e “y\_treino”, que contêm os dados de treinamento, e “x\_teste” e “y\_teste”, que contêm os dados de teste. Esses conjuntos de dados serão utilizados para treinar e avaliar o desempenho do nosso modelo de *machine learning*.

### 3.6.3 DUMMY CLASSIFIER

O *Dummy Classifier* é um tipo de modelo de aprendizado de máquina que faz previsões utilizando estratégias muito simples. Ele não aprende com os dados e faz previsões baseadas em regras simples e heurísticas. Esse tipo de modelo é muito útil para servir como linha de base, permitindo comparar o desempenho de modelos mais complexos.

Treinamos um modelo *Dummy* e fizemos previsões com ele, utilizando os dados de teste. Como resultado, obtivemos um modelo treinado e as previsões feitas por ele, que foram usadas para avaliar o seu desempenho.

Para avaliar o desempenho do modelo, criamos uma matriz de confusão, que nos fornece uma base para calcular algumas métricas de avaliação, como a precisão (*precision*) e a abrangência (*recall*). A precisão foca na exatidão das previsões positivas do modelo, enquanto a abrangência se concentra na quantidade de previsões positivas corretas.

A matriz de confusão do nosso modelo *Dummy* é apresentada na tabela a seguir:

|   | <i>precision</i> | <i>recall</i> | <i>f1-score</i> |
|---|------------------|---------------|-----------------|
| 0 | 0.53             | 0.45          | 0.49            |
| 2 | 0.41             | 0.48          | 0.44            |

Tabela 1: Matriz de confusão simplificada do *Dummy Classifier*.

No nosso caso, a precisão (*precision*) é mais importante, porque uma precisão alta indica que, quando o modelo prevê que uma ação deve ser comprada (ou vendida), é mais provável que essa previsão esteja correta, evitando investir em ações que não irão se valorizar.

Em seguida, exibimos essa matriz de confusão em um gráfico:

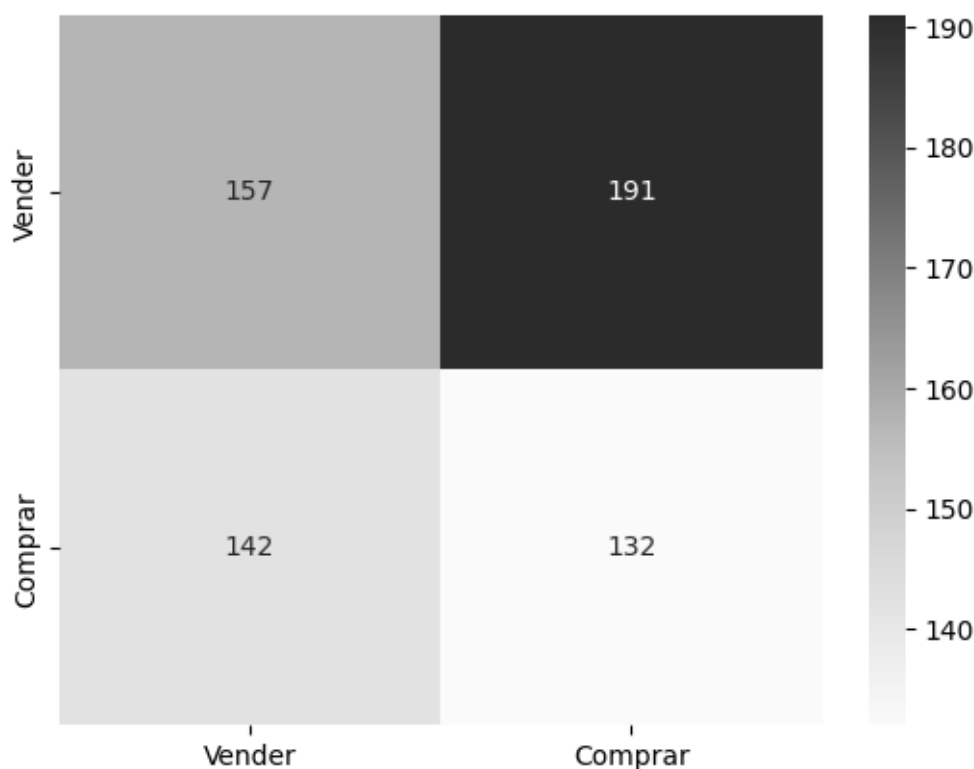


Gráfico 5: Gráfico de calor da matriz de confusão do Dummy Classifier.

No gráfico acima, podemos visualizar as duas classes, que representam os dois resultados possíveis da coluna "Vies" (vender ou comprar).

Os números exibidos indicam as contagens de ocorrências para cada combinação de classe prevista e classe verdadeira. Por exemplo, o modelo previu "Vender" corretamente 157 vezes, ou seja, quando era realmente para vender. Da mesma forma, o modelo previu corretamente "Comprar" 132 vezes, ou seja, quando era realmente para comprar.

#### 3.6.4 MODELOS DE TESTE

Para os objetivos deste trabalho, não há um modelo específico pré-determinado a ser utilizado. Isso porque, no contexto de construção de uma carteira de ações, até mesmo decisões completamente aleatórias podem ser rentáveis e resultar em uma carteira lucrativa.

Dessa forma, selecionamos alguns dos modelos mais utilizados para testá-los, treiná-los e escolher aquele que apresentou o melhor desempenho para prosseguir com o nosso objetivo. Os modelos selecionados para teste foram *AdaBoost*, *Decision Tree*, *RandomForest*,

*ExtraTree*, *Gradient Boost*, *K Nearest Neighbors (KNN)*, *Logistic Regression*, *Naive Bayes*, *Support Vector Machine (SVM)* e Rede Neural.

Cada modelo foi treinado com os dados de treinamento e foram feitas as previsões com os dados de teste. Avaliamos o desempenho de cada modelo, gerando a matriz de confusão e um gráfico de *heatmap* para cada um, a fim de comparar o desempenho e escolher o modelo que melhor se adapta ao nosso problema.

Após os testes, dois modelos se destacaram: *RandomForest* e *SVM*. Ambos apresentaram a mesma precisão (*precision*) de 52%, indicando que 52% dos exemplos classificados como positivos por esses modelos eram de fato positivos. Em relação ao *recall*, o *RandomForest* obteve 57%, enquanto o SVM alcançou apenas 24%. Isso significa que o *RandomForest* foi capaz de identificar uma proporção maior de instâncias positivas reais em comparação com o SVM.

|   | <i>precision</i> | <i>recall</i> | <i>f1-score</i> |
|---|------------------|---------------|-----------------|
| 0 | 0.64             | 0.59          | 0.61            |
| 2 | 0.52             | 0.57          | 0.55            |

Tabela 2: Matriz de confusão simplificada do *RandomForest*.

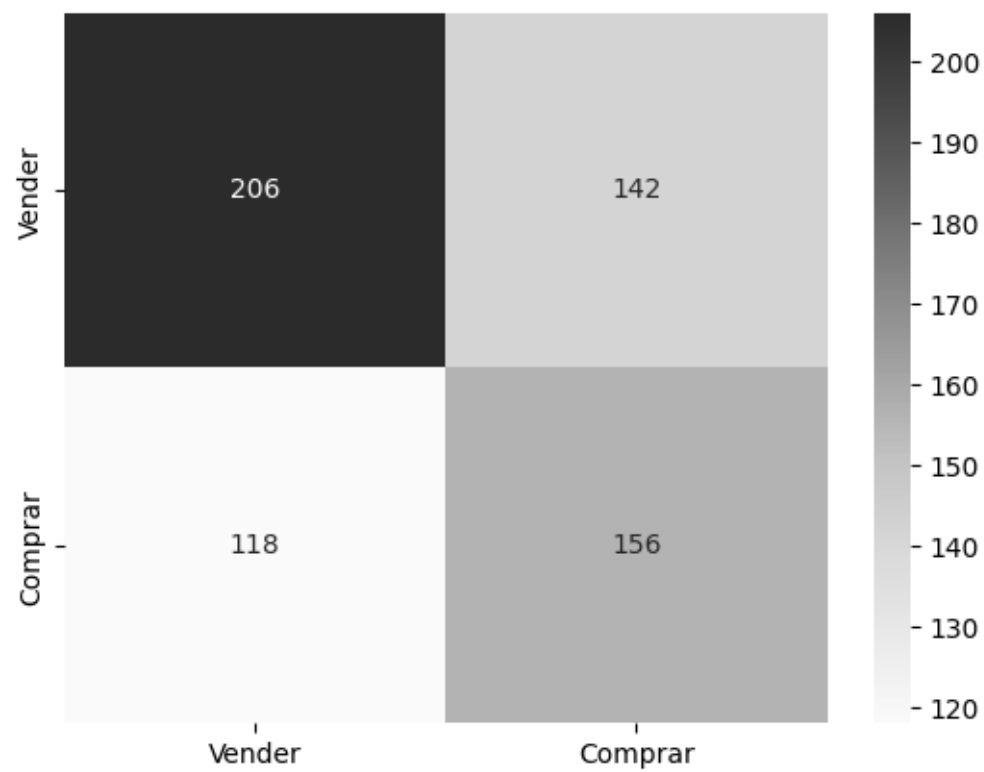


Gráfico 6: Gráfico de calor da matriz de confusão do RandomForest.

|   | <i>precision</i> | <i>recall</i> | <i>f1-score</i> |
|---|------------------|---------------|-----------------|
| 0 | 0.58             | 0.82          | 0.68            |
| 2 | 0.52             | 0.24          | 0.33            |

Tabela 3: Matriz de confusão simplificada do SVM.



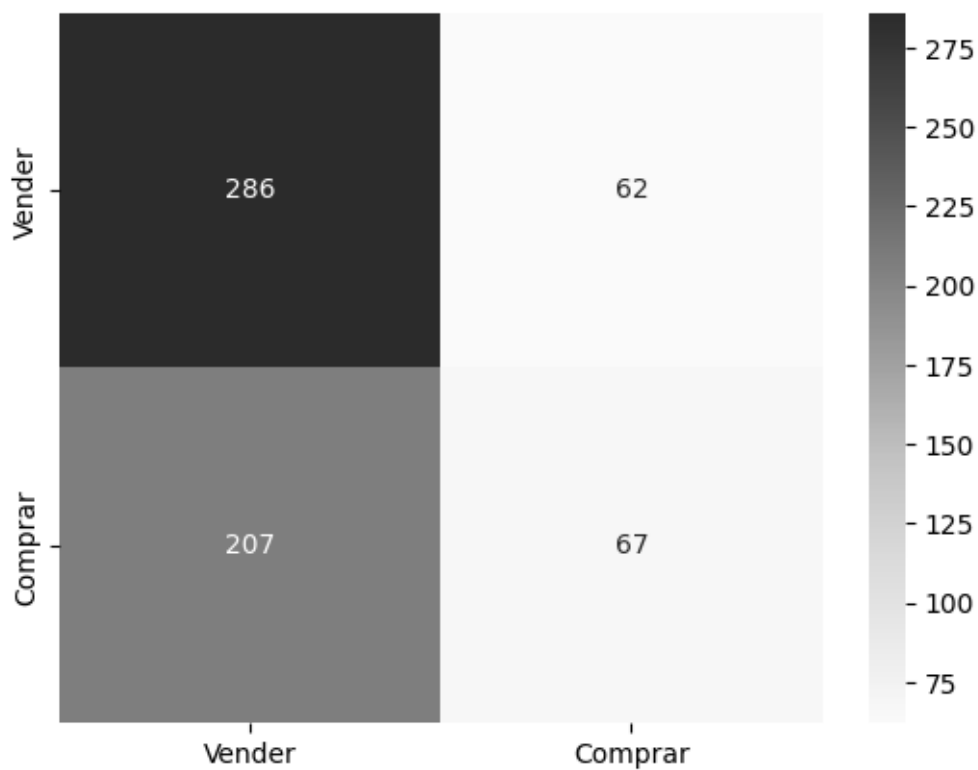


Gráfico 7: Gráfico da matriz de confusão do SVM.

Para este projeto, a precisão é a métrica principal para comparação, enquanto o *recall* é uma métrica secundária. Portanto, o *RandomForest* se torna a escolha mais adequada para o projeto, uma vez que apresenta desempenho igual ao do *SVM* em precisão, mas supera-o em *recall*.

Embora o F1-Score seja importante quando precisão e *recall* têm a mesma importância, esse não é o caso neste projeto, tornando-o menos relevante para a comparação. No entanto, vale destacar que o *RandomForest* também tem um *F1-Score* mais alto, o que reforça sua superioridade em relação ao *SVM*.

### 3.6.5 AJUSTE DE PARÂMETROS

A etapa de Ajuste de Parâmetros, ou *Tuning* do Modelo, é o processo de ajustar e otimizar os parâmetros de um modelo de *machine learning* para melhorar o seu desempenho. Os parâmetros de um modelo são os aspectos internos do modelo que são aprendidos durante o treinamento.

Para otimizar os parâmetros do modelo de melhor desempenho, o *RandomForest*, foi utilizada a técnica de busca em grade (*GridSearchCV*). Essa técnica consiste na definição de uma "grade" de valores de parâmetros e na avaliação de todas as combinações possíveis.

Foram definidas faixas de valores para três parâmetros do modelo: *n\_estimators*, *max\_features* e *min\_samples\_split*. Em seguida, calculou-se a pontuação de precisão para a classe 2 (comprar).

O modelo foi treinado com diferentes combinações de parâmetros e a combinação que maximizou a pontuação de precisão foi escolhida.

Após o processo de ajuste de parâmetros, as previsões foram feitas com os dados de teste, e o desempenho do modelo foi exibido. Uma matriz de confusão e um gráfico *heatmap* foram gerados para avaliar o desempenho do modelo após o ajuste de parâmetros.

|   | <i>precision</i> | <i>recall</i> | <i>f1-score</i> |
|---|------------------|---------------|-----------------|
| 0 | 0.58             | 0.82          | 0.68            |
| 2 | 0.52             | 0.24          | 0.33            |

Tabela 4: Matriz de confusão simplificada do modelo ajustado.

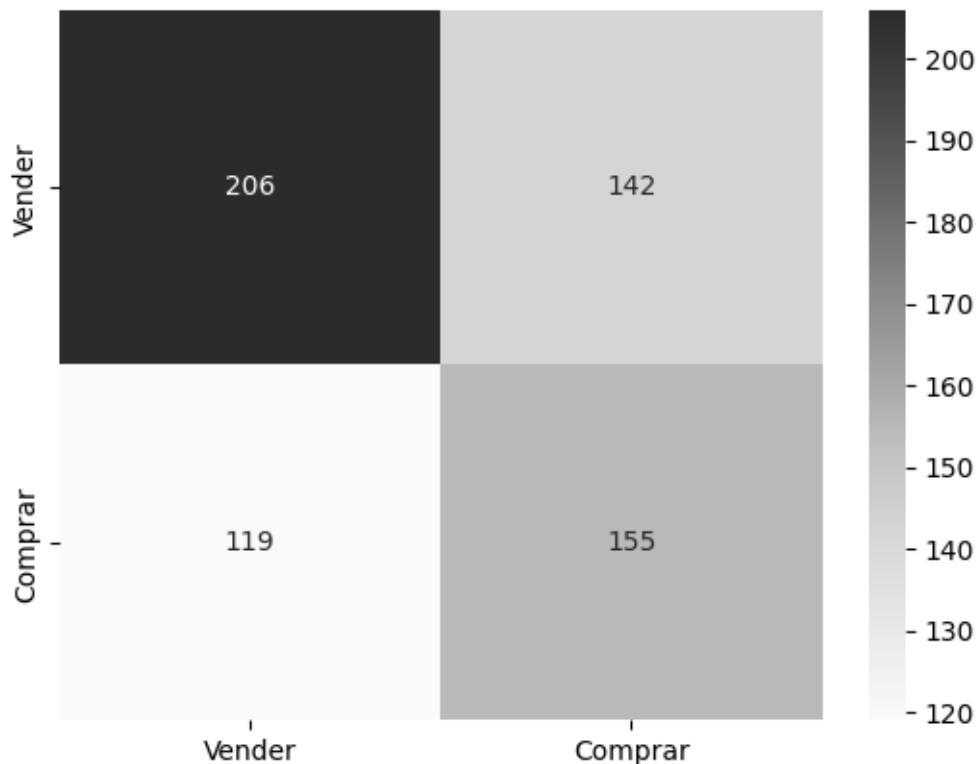


Gráfico 8: Gráfico de calor da matriz de confusão do modelo ajustado.

Comparando com o modelo inicial, sem ajuste de parâmetros, as métricas foram bastante próximas. No entanto, há algumas diferenças sutis. A métrica “*precision*” para a classe 0 é um pouco menor no modelo com ajuste de parâmetros (0.63) em comparação com o modelo sem ajuste (0.64). A métrica *F1-Score* para a classe 2 também é um pouco menor no modelo com ajuste de parâmetros (0.54) em comparação com o modelo sem ajuste (0.55).

No geral, parece que o ajuste de parâmetros não teve um impacto significativo na performance do modelo para essas duas classes específicas. Isso pode sugerir que os parâmetros padrão do modelo *RandomForest* já são razoavelmente bons para esse conjunto de dados.

### 3.7 AVALIAÇÃO

Finalmente, chegamos à etapa de avaliação do nosso modelo de IA. O objetivo nessa fase é verificar se o modelo é capaz de escolher boas ações e auxiliar na construção de uma carteira rentável. Para realizar essa avaliação, utilizaremos o último trimestre dos nossos dados, que não foi utilizado durante as etapas de treinamento e teste. Vale lembrar que o último trimestre dos nossos dados foi descartado porque não tínhamos os dados de cotação e Ibovespa no trimestre seguinte. Isso nos permitirá avaliar o desempenho do modelo de forma mais precisa.

#### 3.7.1 AVALIAÇÃO DA INTELIGÊNCIA ARTIFICIAL

Para realizarmos a avaliação final, executamos os mesmos passos que realizamos anteriormente com a nossa base de dados, mas nos concentrando apenas no último trimestre de cada empresa. Para isso, foi criado um Data Frame específico para armazenar os dados referentes a esse último trimestre.

Em seguida, foram feitas as previsões utilizando o modelo ajustado, utilizando o Data Frame recém-criado com os dados do último trimestre.

Duas carteiras foram criadas: uma carteira inicial e uma carteira final. Se a previsão para uma empresa indicava compra, era adicionado R\$ 1000 à carteira inicial. Em seguida, o percentual de variação entre a cotação final e a cotação inicial era calculado, e o valor final era calculado com base na cotação de R\$ 1000.

Por fim, o saldo inicial foi calculado somando os valores da carteira inicial, e o saldo final foi calculado somando os valores da carteira final.

## 4 RESULTADOS

O resultado da nossa avaliação mostra que, se tivéssemos seguido as recomendações do modelo para comprar ações das empresas listadas, e investido R\$ 1.000 reais em cada uma delas, nossa carteira teria sofrido uma diminuição de valor.

O saldo inicial de R\$ 40.000 representa o total do dinheiro investido, que é a soma de R\$ 1.000 para cada uma das 40 empresas recomendadas para compra pelas previsões do modelo.

O saldo final de R\$ 38.143,85 é a soma dos valores finais das ações dessas 40 empresas no final do trimestre seguinte. Cada valor final é calculado como o produto de R\$ 1000 e a variação percentual nas cotações da ação entre o início e o final do trimestre.

A variação da carteira é calculada como o saldo final dividido pelo saldo inicial, que é 0,95 ou 95%. Isso significa que o valor da nossa carteira de ações diminuiu em 5% ao longo do trimestre.

Uma rentabilidade percentual negativa em relação ao desempenho de uma carteira de ações não necessariamente significa que houve prejuízo. Isso ocorre porque os rendimentos gerados pelas ações podem ter três efeitos diferentes: caso a carteira tenha um desempenho ruim, os rendimentos podem ser insuficientes para compensar as perdas, podem igualar o prejuízo ou até mesmo resultar em ganhos líquidos.

Além disso, uma rentabilidade negativa não significa obrigatoriamente que o desempenho foi inferior ao do mercado em geral. Mesmo que a rentabilidade seja negativa, a carteira ainda pode ter tido um desempenho melhor do que a média do mercado.

### 4.1 AVALIAÇÃO COMPARATIVA

Para avaliar o desempenho de uma carteira de ações, podemos compará-la com um benchmark, que no caso do mercado brasileiro, é frequentemente o Ibovespa. O Ibovespa é um índice que representa o desempenho médio do mercado de ações brasileiro, e nos ajuda a entender se nossa carteira está se saindo melhor ou pior que o mercado em geral. Se o índice Ibovespa tiver um desempenho pior do que a nossa carteira, isso pode ser positivo, indicando que, mesmo em um período de baixa do mercado, nossa carteira foi menos afetada por influências externas do que a média do mercado. Por outro lado, se a nossa carteira ficar atrás do índice, pode ser um sinal de que precisamos reavaliar e ajustar a estratégia.

Para fazer essa comparação, calculamos a variação percentual do índice Ibovespa no último trimestre. O resultado da variação do Ibovespa foi de 0,99%, indicando que o índice teve uma pequena diminuição entre o final do último trimestre e o final do trimestre seguinte.

A carteira de ações recomendadas pelo modelo teve um desempenho pior do que o Ibovespa. A carteira teve uma diminuição de 5%, enquanto o Ibovespa teve uma diminuição de menos de 1%. Isso sugere que, para este período específico, teria sido mais vantajoso investir no Ibovespa do que nas ações específicas recomendadas pelo modelo.

## 5 CONCLUSÃO

Como demonstrado ao longo do trabalho, é cabível enfatizar a relevância do tema em questão, uma vez que este pode exercer um impacto significativo no mercado de ações, especialmente através da análise fundamentalista. A revolução proporcionada pela IA permite analisar grandes volumes de dados, impactando tanto as avaliações de empresas quanto nas decisões financeiras. A utilização da IA na tomada de decisão, com base na análise fundamentalista, é de crescente importância no mercado financeiro, pois contribui para uma tomada de decisão mais informada, rápida e estratégica, além de reduzir erros humanos e avaliar riscos. Portanto, compreender a relação entre IA e análise fundamentalista é fundamental para explorar o seu potencial na otimização dos processos de decisão.

Como principais resultados desta pesquisa, é possível concluir que o modelo de IA não foi bem-sucedido em escolher ações que resultassem em uma carteira rentável, com uma redução de 5% no valor da carteira ao longo do trimestre. Apesar disso, é importante ressaltar que a rentabilidade negativa nem sempre indica prejuízo ou desempenho inferior ao do mercado. No caso deste trimestre, a carteira teve um desempenho inferior ao do IBOV no mesmo período, mas isso não descarta a possibilidade de haver ganhos líquidos em outros trimestres. Portanto, embora o modelo não conseguiu construir uma carteira rentável, esses resultados fornecem percepções importantes sobre os desafios da utilização de modelos de IA para tomada de decisão no mercado de ações.

Para futuras pesquisas acerca do tema, sugerem-se abordagens que considerem a análise de um conjunto de dados maior e mais diversificado, incluindo dados de diferentes períodos. Também seria interessante explorar técnicas de otimização de parâmetros e experimentar diferentes algoritmos de aprendizado de máquina para aprimorar o desempenho do modelo. Além disso, seria relevante identificar e selecionar as variáveis mais relevantes para o modelo de acordo com a abordagem adotada sobre análise fundamentalista ou de acordo com a estratégia pré-estabelecida. Por fim, seria útil realizar uma análise mais aprofundada sobre a rentabilidade da carteira em diferentes cenários de mercado e comparar os resultados com benchmarks do mercado, como o IBOV, para ter uma compreensão mais clara do desempenho do modelo em relação ao mercado em geral.

## REFERÊNCIAS

- ALVES, Vinícius. **Inteligência artificial (IA) pode avaliar riscos e trazer insights na hora de investir**: veja como. Suno. 2023. Disponível em: <https://www.suno.com.br/noticias/inteligencia-artificial-ia-avaliar-riscos-investimentos-va/>. Acesso em: 10 nov. 2023.
- BRAGA, Gustavo Dornelles. **Seleção e análise de uma carteira de ações através do uso da análise fundamentalista**. 2009 Monografia (Administração) - Universidade Federal do Rio Grande do Sul, 2009.
- BRASIL. Casa Civil. Lei n. 6.385, de 06 de dezembro de 1976. **Diário Oficial da União**: Seção 01, Brasília, 07 de dezembro de 1976, ano 155º da Independência e 88º da República. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/l6385.htm](https://www.planalto.gov.br/ccivil_03/leis/l6385.htm). Acesso em: 3 nov. 2023.
- CARVALHO, Valter Pereira de. **Previsão de séries temporais no mercado financeiro de ações com o uso de rede neural artificial**. 2028 Dissertação (Pós-graduação em Engenharia Elétrica e Computação) - Universidade Presbiteriana Mackenzie.
- CAVALCANTE, Francisco; MISUMI, Jorge Yoshio; RUDGE, Luiz Fernando. **Mercado de Capitais**: o que é, como funciona. Elsevier, 2005. 371 p.
- CHOLLET, Francois. **Deep Learning with Python**. Simon and Schuster, v. 2, f. 299, 2017. 597 p.
- COELHO, Filipe Ferreira. **Machine learning e análise técnica como ferramentas para construção de portfólios de renda variável no mercado brasileiro**. 2020 Dissertação (Mestrado Profissional em Economia e Finanças) - Fundação Getúlio Vargas, 2020.
- GONÇALVES, Fabrício. **Como a IA poderá transformar os investimentos**. Capital Aberto. 2023. Disponível em: <https://capitalaberto.com.br/secoes/artigos/como-a-ia-podera-transformar-os-investimentos>. Acesso em: 10 nov. 2023.



INSTITUTO EDUCACIONAL BM&FBOVESPA. **Guia prático de uma das maiores bolsas de valores e derivativos do mundo**. 2017. Disponível em: <https://bvmf.bmfbovespa.com.br/pt-br/download/LivroPQO.pdf>. Acesso em: 10 nov. 2023.

LOBO, Luiz Carlos. Inteligência artificial, o Futuro da Medicina e a Educação Médica. **Revista Brasileira de Educação Médica**, v. 42, n. 3, p. 3-8, jul. 2018.

PINHEIRO, Juliano Lima. **Mercado de Capitais**: fundamentos e técnicas. 4 ed. Atlas, 2008.

REIS, Tiago. **Mercado de ações**: como funciona? Entenda tudo sobre. Suno. Disponível em: <https://www.suno.com.br/guias/mercado-de-acoes/>. Acesso em: 10 nov. 2023.

RUSSELL, Stuart J.; NORVIG, Peter. **Artificial Intelligence**: A Modern Approach. 2 ed. Saddle River, New Jersey 07458: Pearson Education, Inc, 2004.

ÍNDICE Brasil 100 (IBrX 100 B3). B3. Disponível em: [https://www.b3.com.br/pt\\_br/market-data-e-indices/indices/indices-amplos/indice-brasil-100-ibrx-100-composicao-da-carteira.htm](https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/indice-brasil-100-ibrx-100-composicao-da-carteira.htm). Acesso em: 10 nov. 2023.