

Catálogo na publicação
Seção de Catalogação e Classificação

L864s Lopes, André Hugo Ramalho.

Sistema de recomendação de investimento em ativos de renda variável usando técnicas de inteligência artificial / André Hugo Ramalho Lopes. - João Pessoa, 2025.

17 f.

Orientação: Gledson Elias Da Silveira.

Coorientação: Leandro Carlos De Souza.

TCC (Graduação) - UFPB/CI.

1. Sistema de Recomendação. 2. Investimento. 3. Renda variável. 4. Inteligência Artificial. I. Da Silveira, Gledson Elias. II. De Souza, Leandro Carlos. III. Título.

UFPB/CI

CDU 004.8

Sistema de Recomendação de Investimento em Ativos de Renda variável usando Técnicas de Inteligência Artificial

André H. R. Lopes¹, Gledson Elias^{1,2}, Leandro C. de Souza^{1,2}

¹Centro de Informática – Universidade Federal da Paraíba (UFPB)

²Departamento de Informática – Universidade Federal da Paraíba (UFPB)

andre.lopes@academico.ufpb.br, gledson@ci.ufpb.br, leandro@ci.ufpb.br

Abstract. *This paper presents the development of a stock investment recommendation system based on fundamental indicators and artificial intelligence techniques. By collecting financial data on shares from financial indicators of stocks traded on the Brazilian stock exchange, clustering methods are applied to group similar assets. The system allows the creation of a synthetic stock profile based on investor preferences, which is then matched to the most similar real asset using the nearest neighbor approach. The proposed method avoids rigid filters and provides a personalized recommendation based on real data, offering a more flexible and accurate tool for asset selection in the stock market.*

Resumo. *Este trabalho apresenta o desenvolvimento de um sistema de recomendação de investimentos em ativos de renda variável, utilizando indicadores fundamentalistas e técnicas de inteligência artificial. Através da coleta de dados de ações da bolsa de valores brasileira, são aplicados métodos de clusterização para agrupar ativos semelhantes. O sistema permite criar um perfil sintético com base nas preferências do investidor, sendo esse perfil associado ao ativo real mais próximo utilizando a técnica de vizinho mais próximo. A abordagem proposta evita filtros rígidos e oferece recomendações personalizadas baseadas em dados reais, tornando o processo de seleção de ativos mais flexível e preciso.*

1. Introdução

O advento da Inteligência Artificial (IA) revolucionou diversas áreas do conhecimento, oferecendo ferramentas capazes de processar grandes volumes de dados, identificar padrões complexos e propor soluções personalizadas. No contexto financeiro, essa tecnologia surge como um divisor de águas, especialmente em ambientes dinâmicos como o mercado de renda variável, onde a volatilidade e a multiplicidade de variáveis desafiam até os investidores mais experientes. A aplicação de técnicas de IA para investimentos não apenas otimiza a análise de informações, mas também redefine estratégias tradicionais, permitindo abordagens baseadas em evidências e adaptáveis às necessidades individuais [Heaton et al. 2016].

Este trabalho foca na intersecção entre tecnologia e finanças, propondo um sistema de recomendação que utiliza métodos de aprendizado de máquina para transformar a seleção de ativos em um processo mais preciso e alinhado aos objetivos dos usuários. Nesta direção, o presente trabalho tem como objetivo explorar a aplicação da inteligência artificial no desenvolvimento de um sistema de recomendação de investimento em ativos de renda variável. A proposta se baseia na coleta de dados de indicadores fundamentalistas das ações negociadas na bolsa de valores, seguida pela

utilização de técnicas de clusterização para agrupar as ações em categorias semelhantes. O que torna original o sistema de recomendação proposto é a capacidade de, a partir do agrupamento de ações, realizar uma escolha personalizada, permitindo ao investidor identificar a ação que melhor se alinha com suas metas e estratégias.

A proposta não é simplesmente descartar ações que não atendem a critérios preestabelecidos de forma rígida, mas sim evitar a imposição de filtros fixos que possam limitar a exploração de oportunidades relevantes. O objetivo é destacar, entre os ativos disponíveis no mercado, aqueles que melhor se alinham ao perfil e às preferências do investidor. A exclusão automática de ativos que não se enquadram em determinados padrões pode restringir significativamente o leque de opções, especialmente quando se considera o desafio de ajustar múltiplos indicadores de forma criteriosa para garantir a permanência de ao menos um candidato promissor. Assim, adotar uma abordagem mais flexível e adaptativa pode ampliar as possibilidades de identificação de ativos com alto potencial de investimento.

O processo de recomendação consiste em direcionar o investidor a ativos que não apenas se encaixam em seu perfil, mas que também têm potencial para alcançar seus objetivos financeiros. Para isso, o modelo identifica o grupo ao qual a ação desejada pertence e, em seguida, calcula o vizinho mais próximo do centróide deste grupo. O resultado é uma recomendação mais precisa e fundamentada, baseada em dados reais e em técnicas avançadas de inteligência artificial.

O mundo dos investimentos em ativos de renda variável é um cenário complexo e desafiador, repleto de oportunidades e riscos [Lo 2004]. A busca por estratégias eficazes de investimento é uma tarefa contínua para investidores e profissionais do mercado financeiro. Nesse contexto, a utilização de técnicas de inteligência artificial se revela uma ferramenta poderosa e inovadora, capaz de proporcionar uma abordagem mais precisa e personalizada na tomada de decisões de investimento.

1.1. Definição do Problema

Os investidores em ativos de renda variável enfrentam o desafio de identificar as melhores oportunidades de investimento em meio a uma vasta quantidade de informações, como os diversos indicadores financeiros, que denotam a "saúde" de um negócio. A complexidade, volatilidade e a dinâmica do mercado dificultam a tarefa de selecionar ações que se alinhem com as metas, expectativas e estratégias individuais de cada investidor. Este trabalho pretende desenvolver um sistema de recomendação que utilize técnicas de inteligência artificial, como agrupamento e k-vizinhos mais próximos, para resolver esse problema. A questão central que motiva esta pesquisa é: Como podemos aplicar técnicas de inteligência artificial para criar um sistema de recomendação que identifique ações alinhadas com os objetivos, expectativas e estratégias dos investidores de maneira eficiente e personalizada?

1.2. Premissas e Hipóteses

A principal premissa deste trabalho é que a aplicação de técnicas de inteligência artificial pode recomendar investimentos em ativos de renda variável de acordo com as necessidades, objetivos e expectativas de um investidor. Parte-se da hipótese de que o

agrupamento de ações com base em valores de indicadores fundamentalistas permitirá a criação de grupos "homogêneos", nos quais é possível identificar ativos que possuem comportamento semelhante, em relação aos seus indicadores. Além disso, assume-se que a análise de proximidade dentro desses grupos pode fornecer recomendações mais precisas e úteis para os investidores [PORTAL DATA SCIENCE, 2023].

A hipótese geral é que um sistema de recomendação que leve em consideração as características esperadas pelo investidor, em cada indicador financeiro, pode resultar na recomendação de um ativo que se adeque aos interesses reais do investidor sem excluir opções, mas identificando um ativo real que mais se alinha com a pretensão do usuário.

1.3. Objetivos

O objetivo geral deste trabalho é desenvolver um sistema de recomendação de investimento em ativos de renda variável, utilizando técnicas de inteligência artificial. Esse sistema busca proporcionar uma abordagem mais precisa e personalizada na seleção de ações, auxiliando os investidores na identificação de oportunidades que estejam alinhadas com seus objetivos financeiros.

Para alcançar esse objetivo, é fundamental coletar, analisar e tratar uma vasta quantidade de dados de indicadores fundamentalistas das ações negociadas na bolsa de valores. Esses indicadores são essenciais para entender a saúde financeira e o potencial de crescimento das empresas. A análise desses dados permitirá a construção de um banco de informações robusto, que servirá de base para as etapas subsequentes do projeto.

Com os dados em mãos, o próximo passo é aplicar técnicas de agrupamento. O agrupamento é um método de aprendizado de máquina que agrupa objetos similares em classes. No contexto deste trabalho, essa técnica será utilizada para agrupar ações que compartilhem características fundamentais semelhantes. Esse processo envolve a escolha de um algoritmo de agrupamento apropriado e a determinação do número ideal de grupos, de modo a garantir que os grupos formados sejam homogêneos e compartilham semelhanças em relação aos valores de seus indicadores.

Uma vez estabelecidos os grupos, o usuário já pode definir suas preferências em relação a cada indicador fundamentalista para definir um perfil. Com base nesse perfil, o sistema localizará o grupo de ações mais adequado e, dentro desse grupo, utilizará uma técnica de proximidade (k-vizinhos mais próximos) para encontrar a ação que mais se aproxima do centróide do grupo. Essa abordagem permite que o sistema ofereça recomendações personalizadas e fundamentadas em dados reais, ao invés de seguir um filtro fixo que possa limitar as opções do investidor.

Por fim, este trabalho também busca contribuir para o desenvolvimento do campo de sistemas de recomendação de investimentos, oferecendo uma abordagem inovadora que combina técnicas de inteligência artificial com a análise de dados financeiros. A crescente disponibilidade de dados e o avanço das tecnologias de processamento e análise possibilitam a criação de sistemas cada vez mais sofisticados e personalizados. O sistema proposto tem o potencial de não apenas melhorar a tomada de

decisões dos investidores, mas também de avançar o conhecimento na interseção entre inteligência artificial e finanças.

2. Conceitos gerais e trabalhos relacionados

A fim de embasar a proposta deste trabalho, é fundamental apresentar os conceitos teóricos e as abordagens metodológicas que sustentam o uso da inteligência artificial na recomendação de ativos financeiros. Para isso, esta seção aborda, inicialmente, os fundamentos da análise fundamentalista e sua aplicação no contexto da renda variável. Em seguida, discute-se o uso de dados financeiros na modelagem estatística, destacando técnicas como clusterização e K-vizinhos mais próximos (KNN), além de examinar ferramentas existentes no mercado que auxiliam investidores na escolha de ações. Por fim, busca-se evidenciar as lacunas dessas soluções e como o uso de técnicas inteligentes pode oferecer uma abordagem mais flexível e personalizada para os investidores.

2.1. Análise fundamentalistas de ativos de renda variável

A análise fundamentalista é uma metodologia que visa avaliar o valor intrínseco de uma empresa por meio do exame minucioso de suas demonstrações financeiras (como balanços patrimoniais e fluxos de caixa) e indicadores de desempenho como dito por Graham e Dodd (2009). De acordo com os preceitos clássicos de Graham e Dodd (2009), essa metodologia baseia-se na premissa de que o mercado pode subvalorizar ou supervalorizar ativos temporariamente, criando oportunidades de investimento quando há divergência entre o preço de mercado e o valor real calculado.

Damodaran (2007) reforça que esse processo permite identificar discrepâncias entre o preço de mercado e o valor real da empresa, oferecendo bases para decisões de investimento de longo prazo. Além disso, a análise fundamentalista incorpora perspectivas quantitativas que proporcionam uma visão do potencial de crescimento e da resiliência do negócio em diversos cenários, dessa forma, permite-se identificar discrepâncias entre o preço de mercado e o valor real da empresa, oferecendo oportunidades de investimento fundamentadas [Damodaran, 2012].

2.2. Análise de dados de indicadores fundamentalistas de ativos de renda variável

A análise de dados estatísticos ou aprendizado estatístico segundo James, Witten, Hastie e Tibshirani (2013) é um vasto arsenal de ferramentas para se entender os dados. Os dados de indicadores fundamentalistas de uma ação, por exemplo, são fontes de informações importantes acerca de seu preço momentâneo (subvalorizado ou valorizado) conforme Villaschi et al. (2011). Com isso, pode-se tentar conhecer e entender o comportamento de uma determinada ação, influenciando diretamente na tomada de decisão de uma possível compra ou venda do ativo.

Conforme James et al. (2013), a aplicação de modelos e algoritmos de *clustering* permite classificar e agrupar dados de natureza similar por suas variáveis-chave (como ROE, margem líquida e crescimento de receita). Villaschi et al. (2011) demonstram, por exemplo, que indicadores como EV/EBITDA e P/L projetivo são preditores eficazes de subvalorização em setores cíclicos, enquanto métricas de endividamento (Dívida

Líquida/EBITDA) são críticas para avaliar riscos em empresas altamente alavancadas. Com essas informações e interpretações, o investidor pode inferir o comportamento de um determinado ativo no mercado e buscar opções de investimento de acordo com as características desejadas.

2.3. Modelagem estatística para agrupamento de ações

A clusterização, ou agrupamento, é uma técnica estatística de aprendizado de máquina não supervisionada que agrupa dados em subconjuntos homogêneos (clusters) com base em similaridades entre suas características, em concordância com James et al. (2013). O objetivo é identificar padrões intrínsecos aos dados, permitindo que observações semelhantes possam ser categorizadas em grupos coesos [Hastie et al. 2009]. No contexto financeiro, a clusterização pode segmentar empresas da bolsa de valores brasileira (B3) conforme indicadores como ROE (*Return on Equity* ou retorno sobre patrimônio líquido), P/L (preço sobre lucro) e nível de endividamento, revelando setores subvalorizados ou comportamentos de mercado correlacionados.

2.4. Modelagem estatística para reconhecimento de ações semelhantes

O algoritmo K-Vizinhos Mais Próximos (KNN - *K-Nearest Neighbors*) é um algoritmo de aprendizado de máquina supervisionado utilizado para classificação e regressão. Sua lógica central baseia-se no princípio de que dados semelhantes tendem a estar próximos no espaço de características.

O funcionamento do algoritmo *K-Nearest Neighbors* (KNN) baseia-se no princípio da proximidade: um novo dado é classificado de acordo com a maioria dos seus “k” vizinhos mais próximos no espaço de características. A medida de distância mais comum utilizada para calcular a similaridade entre os pontos é a distância euclidiana (Hastie; Tibshirani; Friedman, 2009). No processo de classificação, o K-NN busca os “k” exemplos de treinamento mais próximos da nova instância e atribui a ela a classe mais frequente entre esses vizinhos [James et al., 2013].

Conforme Hastie et al. (2009), o algoritmo KNN se baseia na classificação dos “k” vizinhos mais próximos, a classe com maior representação entre eles será dada ao objeto a ser classificado. Diante disso, dentro de um grupo de ações já classificadas como de mesmo grupo, tem-se a possibilidade de achar aquela que mais represente aquele grupo, ou seja, a mais próxima do centróide.

2.5. Trabalhos Relacionados

Existem algumas ferramentas que auxiliam investidores a encontrarem ações de seu perfil de investimento, que principalmente se baseiam nas ideias de análise de dados de indicadores fundamentalistas, como destacado por Bodie, Kane e Marcus (2014), que afirmam que os investidores frequentemente utilizam filtros fundamentados em métricas financeiras para selecionar ativos que atendam aos seus objetivos de retorno e tolerância a risco.

Adicionalmente, a utilização de plataformas como o Status Invest e o Fundamentus exemplifica como a análise de dados fundamentalistas tem sido incorporada de maneira prática para facilitar a triagem de ações [Status Invest, 2025;

Fundamentus, 2025]. Essas ferramentas facilitam a busca dos investidores por novos ativos com potencial de lucro, além de otimizar o tempo demandado para esta tarefa, já que o mercado de valores brasileiro já conta com milhares de ativos e cerca de 400 empresas listadas [TORO INVESTIMENTOS, 2024].

A visualização clara de indicadores fundamentalistas é essencial para democratizar o acesso a informações estratégicas, especialmente em um mercado complexo como o brasileiro. Plataformas como a Suno Analítica [Suno Analítica, 2025] oferecem painéis detalhados para ações específicas, como a ABEV3, exibindo métricas como P/L, EV/EBITDA, ROIC e histórico de dividendos em gráficos interativos (Figura 1).

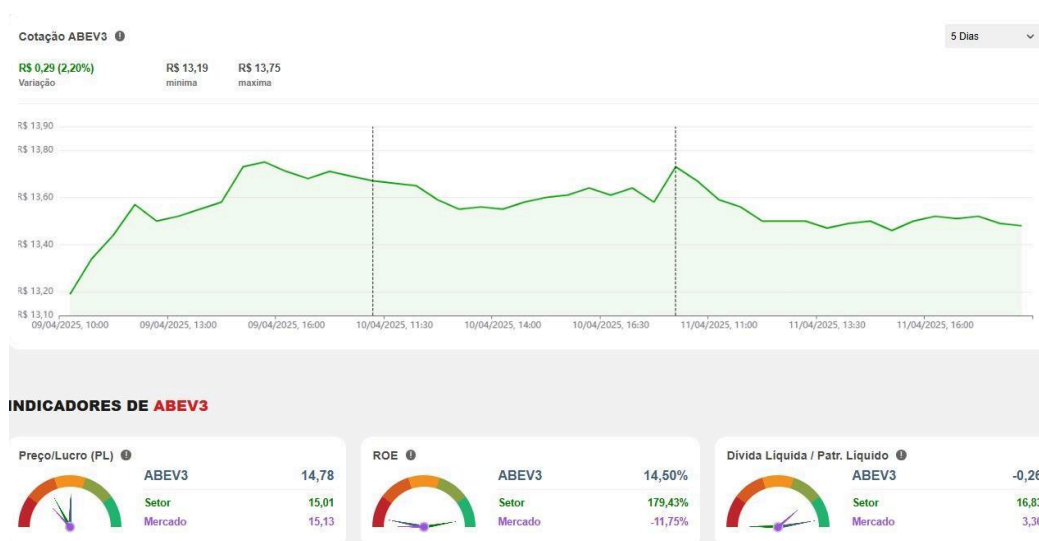


Figura 1. Painel de visualização de indicadores da Suno Analítica

Na Figura 1 é possível verificar informações sobre a faixa de valores dos indicadores entre todos os ativos. Por exemplo, para o indicador Preço/Lucro (PL) pode-se ver 4 informações comparativas. Dentre estas informações estão 2 indicadores relativos ao mercado em geral, onde o melhor valor é representado pelo ponteiro à direita e o pior valor é indicado pelo ponteiro à esquerda. Além do valor do indicador para a própria ação, há também 2 indicadores relativos ao setor da ação e a média do mercado como um todo.

Complementando essa abordagem, ferramentas de busca avançada permitem filtrar empresas listadas na B3 com base em critérios personalizados. Os sites Status Invest e Investidor 10 [Status Invest, 2025; Investidor 10, 2025], por exemplo, possibilitam que investidores selecionem ações por faixas de valores de indicadores como: P/VP, DY mínimo, liquidez média diária ou crescimento de lucros. Essas plataformas automatizam a triagem de ativos, reduzindo horas de análise manual e destacando oportunidades alinhadas ao perfil e objetivos do usuário. As Figuras 2 e 3 apresentam as buscas avançadas dos sites Status Invest e Investidor 10, respectivamente.

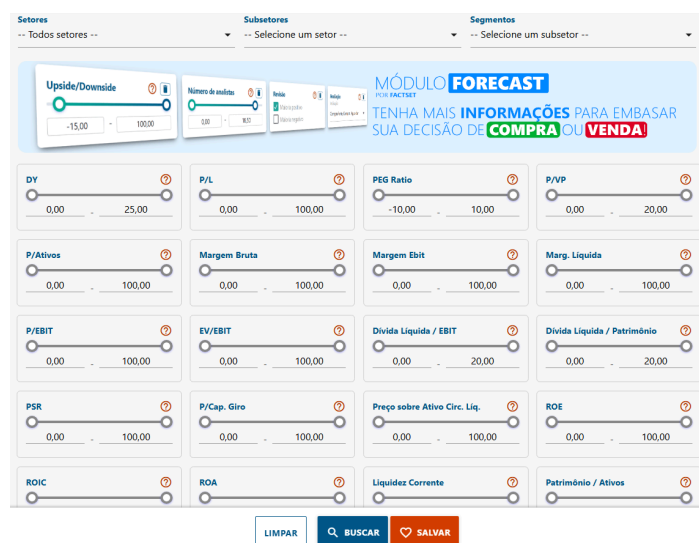


Figura 2. Busca avançada por ações no site Status Invest.

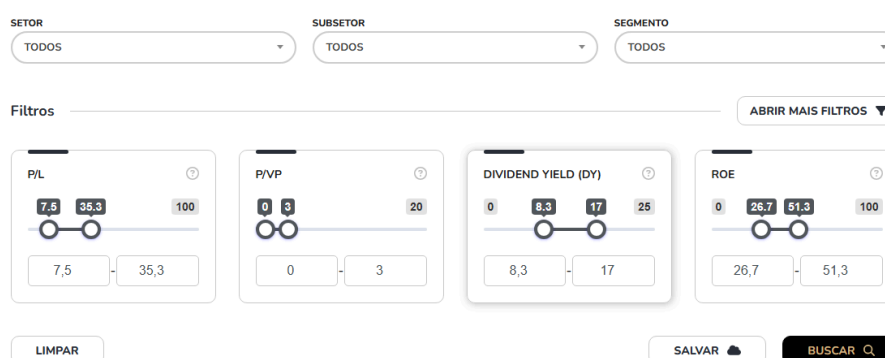


Figura 3. Busca avançada por ações no site Investidor 10.

Além das funcionalidades já descritas, como pode ser observado nas Figuras 2 e 3, pode-se também acrescentar filtros mais específicos por setor, subsetor e segmento da ação. O uso de análises mais robustas se faz presente em versões pagas da plataforma, que contam com previsões, comparações e indicações como expresso no *banner* da Figura 2.

É importante mencionar que nenhuma das aplicações citadas faz menção ao uso de inteligência artificial nessas plataformas. As aplicações fazem filtragem das ações do mercado, baseadas exclusivamente nos valores de seus indicadores fundamentalistas, e mostram como resultado as ações enquadradas. Caso os filtros sejam ajustados de forma peculiar, pode não haver resultados na busca mesmo que o usuário seja mais flexível em sua escolha. Esse impasse traz à tona um outro problema para procura de ações: otimização de valores de filtro para resultados assertivos (encontrar uma ação desejada) levando em conta diversos indicadores.

Vale ressaltar que não faz parte do escopo deste trabalho avaliar os impactos positivos e negativos dos diferentes indicadores fundamentalistas nas respectivas ações,

sendo, portanto, responsabilidade do usuário expressá-los com valores que representam sua experiência e intuição em investimentos financeiros.

3. Metodologia

A condução deste estudo envolveu uma abordagem metodológica robusta e cuidadosamente estruturada, compreendendo as seguintes etapas: coleta de dados, tratamento de dados, modelo de clusterização, modelo de vizinhos mais próximos e aplicação web de recomendação de investimentos.

3.1. Coleta de dados

A etapa inicial deste estudo envolveu a coleta de dados por meio do pacote "fundamentus" [Ferreira, M. 2021], uma ferramenta disponibilizada no gerenciador de pacotes do Python. Este pacote facilita o acesso a dados históricos de indicadores fundamentalistas de ativos negociados na bolsa de valores.

As funções abrangentes presentes no pacote permitem a extração de valores de diversos indicadores em períodos ajustáveis, incluindo preços, valores de dividendos, valores de distribuição de juros sob capital próprio. Além disso, as funções possibilitam a segregação por setores de atuação das ações, como Energia, Telecomunicações, entre outros.

Para este trabalho foram coletados os dados de 19 indicadores fundamentalistas de ações, dentre eles estão: preço sobre lucro (PL), retorno sobre patrimônio (ROE), retorno sobre capital investido (ROIC), *Dividend yield* (DY), margem de liquidez (MRGLIQ), dívida bruta sobre patrimônio (DIVBPATR), preço sobre valor patrimonial (PVP), preço por ativo (PA), preço sobre capital de giro (PCG), preço da ação sobre lucro antes dos juros e impostos (PEBIT), preço por ativo circulante líquido (PACL), valor da empresa sobre lucro antes de juros e impostos (EVEBIT), valor da empresa sobre lucro antes de juros, impostos, depreciações e amortizações (EVEBITIDA), margem de lucro antes de juros e impostos (MRGEBIT), liquidez corrente (LIQC), liquidez nos últimos 2 meses (LIQ2M), patrimônio líquido (PATRILIQ), crescimento da receita líquida nos últimos 5 anos (C5Y) e Valor da ação multiplicado pelo resultado da divisão do número de ações emitidos pela receita líquida (PSR).

Além disso, dados de 968 ativos da bolsa de valores dos mais diversos nichos de atuação foram coletados para este trabalho. Cada ativo teve seu nome, cotação e respectivos indicadores fundamentalistas armazenados para criação de um banco de dados.

3.2. Tratamento de dados

Em um cenário de desenvolvimento de sistemas cada vez mais orientado por dados, a importância do bom processamento de dados torna-se crucial. Um eficiente processamento de dados não apenas agiliza a extração de informações valiosas, mas também assegura a qualidade e integridade dos dados utilizados para análise. Sistemas robustos de processamento de dados têm a capacidade de tomar decisões mais informadas e estratégicas, ganhando uma vantagem competitiva significativa.

Antes de iniciar qualquer análise, o pré-processamento de dados desempenha um papel essencial na preparação adequada das informações. Um bom pré-processamento elimina ruídos, trata valores ausentes e padroniza formatos, permitindo que os modelos de análise e algoritmos extraiam padrões mais significativos e relevantes. Investir em um bom pré-processamento não apenas melhora a precisão dos resultados, mas também economiza recursos ao longo do ciclo de vida do projeto, contribuindo para a eficácia geral das soluções baseadas em dados.

Os dados obtidos, inicialmente no formato decimal brasileiro e apresentados como strings, passaram por um processo de tratamento. Essa fase incluiu a conversão dos dados para um formato manipulável por bibliotecas especializadas em álgebra linear e outros métodos matemáticos. Posteriormente, todos os dados foram armazenados em uma estrutura de dados do tipo *DataFrame*, viabilizando assim a preparação necessária para o treinamento dos modelos, por meio da importação da biblioteca *pandas*.

Trabalhar com valores de ponto flutuante e com vários significados diferentes pode tornar o modelo pouco preciso, para isso uma das opções válidas para contornar o problema é classificar esse valores por faixas de valores.

Pensando neste problema, cada dado de indicador e seu valor intrínseco (I) foram excluídos os *outliers*, além disso, foi feita uma análise de mercado para se descobrir a média de valores do indicador (μ) e seus desvios padrão (σ) para classificar os valores dos indicadores em 4 classes distintas: classe 1 para valores inferiores a média menos desvio padrão; classe 2 para valores maiores ou iguais a média menos o desvio padrão e menores que a média; classe 3 para valores maiores ou iguais a média e menores que a média mais o desvio padrão e classe 4 para valores maiores ou iguais que a média mais o desvio padrão. Pode-se observar as classes de intervalos descritos na Tabela 1.

Tabela 1. Classes e seus intervalos

Classes	Intervalos abrangidos
Classe 1	$I < \mu - \sigma$
Classe 2	$\mu - \sigma \leq I < \mu$
Classe 3	$\mu - \sigma \leq I < \mu + \sigma$
Classe 4	$I \geq \mu + \sigma$

3.3. Modelo de clusterização

A clusterização, ou agrupamento, é uma técnica fundamental na análise de dados que visa identificar padrões naturais e estruturas subjacentes em conjuntos de dados. Essa abordagem consiste em agrupar pontos de dados semelhantes em clusters (grupos), onde

a similaridade é determinada por características específicas. O algoritmo de clusterização avalia a proximidade entre os pontos, atribuindo-os a grupos distintos com base nas características compartilhadas.

O objetivo é maximizar a similaridade intra-cluster e minimizar a similaridade inter-cluster. A clusterização é amplamente utilizada em diversas áreas, como segmentação de mercado, reconhecimento de padrões em imagens, e na organização de grandes volumes de dados, proporcionando uma compreensão mais profunda e estruturada dos dados analisados.

A clusterização das ações foi realizada utilizando um modelo que agrupa ativos teoricamente pertencentes à mesma classe. Os dados de indicadores fundamentalistas da biblioteca "fundamentus" foram empregados como entrada para esse processo de acordo com as necessidades e escolhas do usuário. O usuário deve escolher a classe de cada indicador que ele gostaria de levar em consideração para que o modelo identifique sua classe.

Para definição do número de *clusters* de cada consulta ao sistema, foi utilizado o método Elbow (Figura 4) para determinar a quantidade ideal de classes em que o conjunto de ações deve ser dividido. A lógica por trás desse método é identificar o ponto de inflexão, indicando o número ótimo de classes. Posteriormente, após o treinamento do modelo, tornou-se possível criar um ativo sintético com indicadores desejados e classificá-lo em uma das classes estratificadas pelo método Elbow.

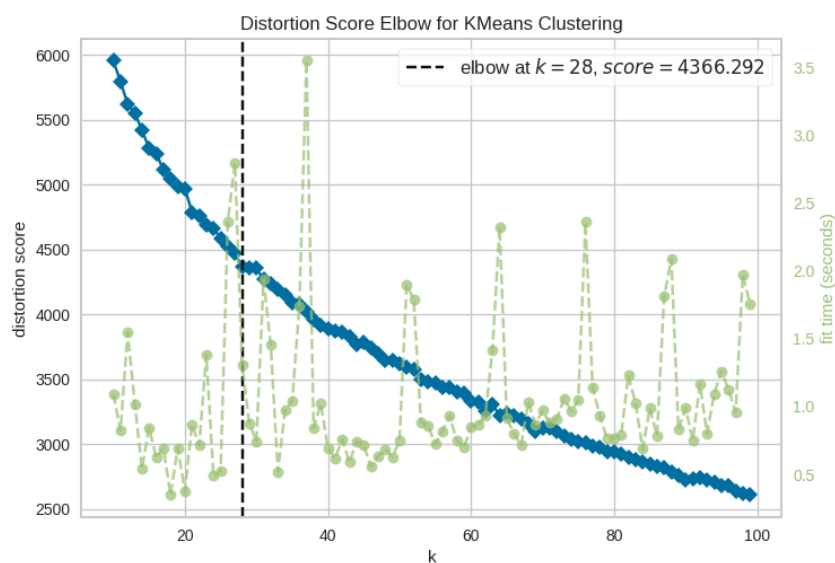


Figura 4. Gráfico de representação do método *Elbow*

O método Elbow (ou método do “cotovelo”) é uma abordagem heurística para definir o número ideal de clusters em algoritmos de agrupamento, como o K-means [Milton et al. 2017]. Funciona calculando-se a soma dos quadrados intra-clusters (WCSS) para diferentes valores de “k” (números de clusters) e plotando esses valores em um gráfico de “k” versus WCSS. À medida que “k” aumenta, o WCSS diminui, pois os dados são divididos em grupos mais compactos. O ponto ótimo é identificado

visualmente como uma "curva de cotovelo", onde a taxa de redução do WCSS muda abruptamente, indicando que adicionar mais clusters não melhora significativamente a compactação dos dados [Milton et al. 2017], como mostrado na Figura 4.

3.4. Modelo de vizinhos mais próximos

O método dos vizinhos mais próximos, ou k-vizinhos mais próximos (k-NN), é um algoritmo de aprendizado de máquina supervisionado utilizado para classificação e regressão. Sua lógica é intuitiva, baseando-se na premissa de que pontos de dados semelhantes tendem a compartilhar características comuns.

No k-NN, a classificação de um novo ponto de dados é determinada pelos votos da maioria dos seus “k” vizinhos mais próximos no espaço de características. No caso deste trabalho adota-se 1 vizinho próximo, pois já temos a classe a qual pertence o objeto.

A medida de proximidade pode ser calculada por diversas métricas, como a distância euclidiana [James et al., 2013] utilizada neste trabalho. A flexibilidade do k-NN torna-o útil em situações em que a relação entre os pontos de dados é complexa e não linear, como no caso deste trabalho que se trata de uma ação com até 19 dimensões (indicadores).

A implementação do modelo de vizinhos mais próximos baseou-se em calcular o mais próximo aos centróides das classes geradas pelo modelo de clusterização através do cálculo da distância euclidiana entre um ativo real e o centróide de uma classe.

Quando um ativo sintético é predito pelo modelo, ele é atribuído a uma classe específica. Utilizando o centróide dessa classe como referência, o modelo de vizinhos mais próximos busca um ativo real da bolsa de valores que compartilha características semelhantes com a classe do ativo sintético previamente criado, ou seja, o ativo com menor distância para o centróide da classe predita.

Essa abordagem permite uma recomendação alinhada com as características do perfil de investimento do usuário, proporcionando um método inovador e eficiente de seleção de ativos na bolsa de valores.

3.5. Aplicação web de recomendação de investimentos

Para visualização de resultados e uma experiência de interação do usuário foi necessária a criação de uma aplicação web para comportar toda a lógica descrita neste trabalho. Toda a lógica de seleção de número de grupos, agrupamento, vizinhos mais próximos e seleção do ativo real ficou encapsulada em uma API REST (interface de programação de aplicações que segue o padrão de arquitetura REST que em português significa “transferência de estado representacional”) que fornecerá todos os dados necessários para consumo. Além disso, foram utilizadas ferramentas para criação de telas web para visualização do conteúdo gerado.

Para a parte de algoritmos lógicos da ideia desenvolvida na metodologia deste trabalho, foi criado um software *backend* em formato de API REST utilizando a

linguagem python e o *framework* Flask. Já para a visualização e interação com a aplicação, foi utilizada a linguagem javascript e o *framework* React.

As classes numeradas são pouco didáticas em significado, diante disso, foi pensado em uma abstração para nomear as classes a fim de simplificar a experiência do usuário em contato com a aplicação. A Tabela 2 mostra as classes, suas abstrações nominais e os intervalos em que o indicador é compreendido no mercado.

Tabela 2. Classes, abstrações nominais e seus intervalos

Classes	Intervalos abrangidos	Abstrações nominais
Classe 1	$I < \mu - \sigma$	Ruim
Classe 2	$\mu - \sigma \leq I < \mu$	Regular
Classe 3	$\mu - \sigma \leq I < \mu + \sigma$	Bom
Classe 4	$I \geq \mu + \sigma$	Ótimo

Dentre as funcionalidades da aplicação, há uma tela de inspeção de ações, na qual qualquer ativo da bolsa de valores pode ser visualizado com seus dados de indicadores fundamentalistas. Este ativo será apresentado ao usuário com os dados dos indicadores nominalmente abstraídos, como mostrado em um exemplo com o ativo ABEV3 na Figura 5.

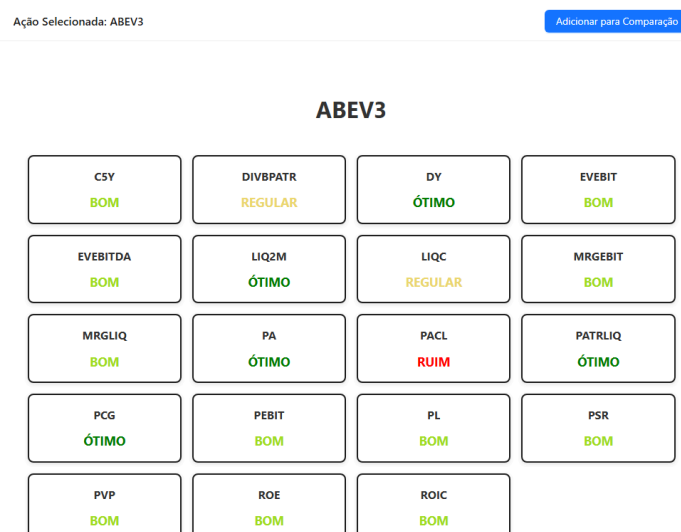


Figura 5. Tela de visualização de ativos

Além disso, na Figura 5, pode-se ver um botão escrito “Adicionar para Comparação”, este botão pode ser utilizado para se visualizar, na mesma tela,

graficamente uma ou mais ações com a finalidade de compará-las. Através de um gráfico de radar, compara-se indicadores como exemplificado na imagem Figura 6, que ilustra a comparação entre os ativos ABEV3, BBDC4 e LAME4.

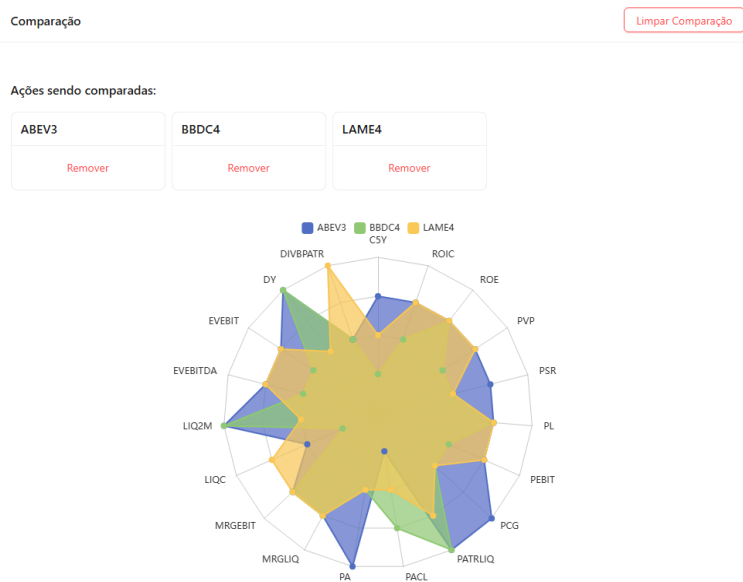


Figura 6. Comparação dos ativos ABEV3, BBDC4 e LAME4

Outra funcionalidade da aplicação, que representa a principal funcionalidade, é a recomendação de ativos com base nas preferências do usuário em relação aos indicadores fundamentalistas do mercado financeiro. Tal ferramenta funciona com a entrada do usuário e suas expectativas para cada indicador financeiro. Assim, o usuário seleciona a classe abstrata para aquele determinado indicador como visto na Figura 7.

Nome da ação

PL 4 - Ótimo ÓTIMO	PVP Seleccione Seleccione 1 - Ruim 2 - Regular 3 - Bom 4 - Ótimo Seleccione	PSR Seleccione	DY 3 - Bom BOM	PA Seleccione
PCG Seleccione	PACL Seleccione	EVEBIT Seleccione	EVEBITDA Seleccione	
MRGEBIT Seleccione	MRGLIQ Seleccione	ROIC Seleccione	ROE Seleccione	LIQC Seleccione
LIQ2M Seleccione	PATRLIQ Seleccione	DIVBPATR Seleccione	CSY Seleccione	

[Encontrar Ação Similar](#)

Figura 7. Seleção de classes dos indicadores

Após a seleção das classes de indicadores, como indicado na Figura 8, pode-se iniciar o processo de “Encontrar Ação Similar”, no qual haverá a atuação da lógica

descrita neste trabalho para, literalmente, encontrar a ação real que mais se assemelha à ação desejada. O resultado de uma dessas buscas pode ser visto na Figura 9.

Minha Ação

PL 4 - Ótimo ÓTIMO	PVP 3 - Bom BOM	PSR Selecione	DY 3 - Bom BOM	PA Selecione
PCG Selecione	PEBIT Selecione	PACL Selecione	EVEBIT Selecione	EVEBITDA 4 - Ótimo ÓTIMO
MRGEBIT Selecione	MRGLIQ Selecione	ROIC 3 - Bom BOM	ROE 3 - Bom BOM	LIQC Selecione
LIQ2M Selecione	PATRLIQ 3 - Bom BOM	DIVBPATR Selecione	CSY Selecione	

[Encontrar Ação Similar](#)

Figura 8. Indicadores pretendidos selecionados

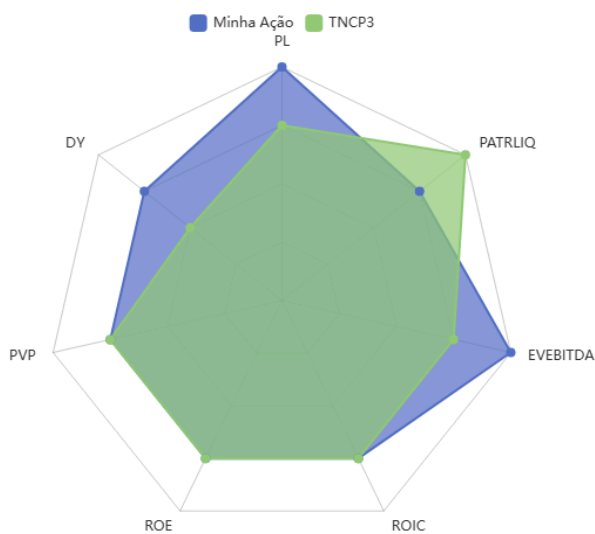


Figura 9. Resultado indicado, a ação encontrada e a ação criada pelo usuário.

4. Resultados e Discussões

A aplicação web desenvolvida neste trabalho avança significativamente em relação às ferramentas tradicionais de seleção de ativos ao combinar técnicas de inteligência artificial com uma interface intuitiva e visualmente rica. Plataformas como Status Invest e Investidor 10 baseiam-se em filtros numéricos estáticos, que frequentemente resultam em buscas sem resultados quando os critérios são muito específicos.

Em contraste, o sistema proposto permite ao usuário definir um perfil sintético de ativo (Figura 7), selecionando preferências qualitativas ("Ótimo" para ROE ou

"Bom" para *Dividend Yield*), sem exigir conhecimento técnico sobre intervalos numéricos. Essa abordagem elimina a frustração de ajustes manuais repetitivos e amplia as opções disponíveis, garantindo que sempre haja uma recomendação alinhada às expectativas do investidor.

A interface da aplicação destaca-se pela integração de visualizações que simplificam a análise complexa de indicadores. Na tela de inspeção de ativos (Figura 5), por exemplo, cada indicador é classificado nominalmente ("Ruim", "Regular", "Bom", "Ótimo"), substituindo a apresentação crua de valores numéricos por uma linguagem acessível. Essa abstração é particularmente útil para investidores leigos, que podem compreender rapidamente o desempenho de um ativo sem precisar interpretar métricas técnicas.

Além disso, a funcionalidade de comparação via gráfico de radar (Figura 6) permite visualizar múltiplos ativos simultaneamente, identificando pontos fortes e fracos em indicadores como P/L e ROIC. Essa capacidade de análise comparativa é ausente em ferramentas como Suno Analítica [Suno Analítica, 2025], que focam em painéis individuais, limitando a visão estratégica do usuário.

O cerne da inovação reside no mecanismo de recomendação, ilustrado nas Figuras 8 e 9. Ao criar um perfil sintético (Figura 8), o sistema identifica o ativo real mais próximo (TNCP3 na Figura 9) utilizando distância euclidiana em relação ao centróide do cluster gerado. Em uma nova comparação (Figura 10), com indicadores diferentes da Figura 9 (adicionado o indicador DIVBPATR classificado como "Bom") mostra-se como mais próximo outro ativo (CPLE11). Dessa forma, evidencia-se que os indicadores do ativo real recomendado alinham-se às preferências do usuário, levando em consideração as mudanças em cada indicador e sua classificação.

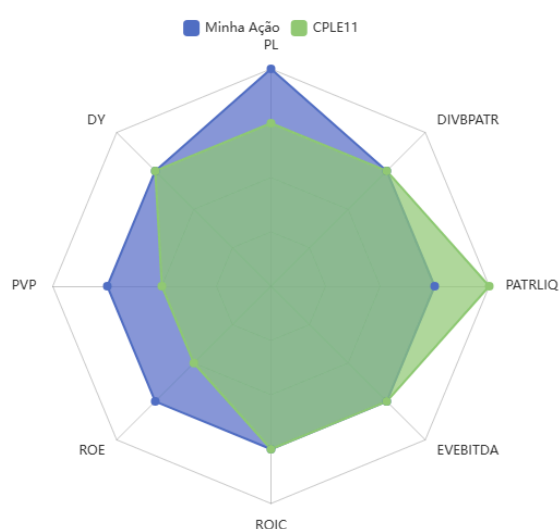


Figura 10. Comparação de ação criada e ação real indicada

Apesar das vantagens, os desafios persistem. A classificação em quatro categorias nominais, embora simplifique a experiência, pode mascarar variações sutis dentro de uma mesma classe. Por exemplo, dois ativos classificados como "Bom" em

P/L podem estar em extremos opostos do intervalo, impactando sua adequação a estratégias de longo prazo.

Em síntese, a aplicação demonstra que a integração de IA em sistemas de recomendação de ativos pode ser uma opção viável e simplificada em comparação a ferramentas atuais. Ao substituir filtros rígidos por recomendações personalizadas e visualizações claras, o sistema torna o processo de seleção de ativos mais simples, democrático e alinhado às preferências reais dos investidores.

5. Conclusão

Este trabalho propôs um sistema de recomendação de investimentos em ativos de renda variável que integra técnicas de inteligência artificial, como clusterização e k-vizinhos mais próximos, com uma interface web intuitiva, visualmente orientada e facilitada para os usuários.

Ao substituir filtros rígidos por uma abordagem baseada em perfis sintéticos, o sistema oferece recomendações personalizadas que se adaptam às preferências dinâmicas dos investidores, resolvendo um problema comum em ferramentas tradicionais: a ausência de resultados devido a critérios excessivamente restritivos.

Além disso, a aplicação desenvolvida não apenas simplifica a interpretação de indicadores complexos por meio de classificações nominais ("Ruim" a "Ótimo"), mas também introduz funcionalidades inovadoras, como a comparação gráfica via radar e a transparência na correspondência entre perfis desejados e ativos reais.

O caso de uso do sistema como possível ferramenta de interesse do investidor, recomendando ativos reais mais próximos do pretendido, reforça sua utilidade prática em cenários reais. No entanto, o trabalho também evidencia desafios, como a necessidade de atualização contínua de dados e a simplificação excessiva inerente à classificação nominal, que merecem atenção em pesquisas futuras.

Para um possível futuro trabalho, pode-se investigar a utilização de mais classes de distinção do mercado, a fim de distinguir mais precisamente os dados dos indicadores. Além disso, o sistema pode ser atualizado para mostrar não só um, mas outros vizinhos mais próximos como forma de opção de investimento. Em suma, o trabalho abre caminho para sistemas de recomendação mais inteligentes, flexíveis e alinhados às demandas do mercado financeiro moderno.

7. Referências

- JAMES, G. et al. *An Introduction to Statistical Learning with Applications in R*. New York: Springer, 2013.
- LO, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*.
- HEATON, J. B.; POLSON, N. G.; WITTE, J. H. (2016). Deep learning in finance. arXiv preprint arXiv:1602.06561.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". 2. ed. New York: Springer, 2009.

- GRAHAM, Benjamin; DODD, David L. Security Analysis: Principles and Techniques. 6. ed. New York: McGraw-Hill, 2009.
- DAMODARAN, A. (2007). Strategic Risk Taking: A Framework for Risk Management. Wharton School Publishing.
- DAMODARAN, A. Investment Valuation: Tools and Techniques for Determining the Value of Any Asset. Wiley, 2012.
- BODIE, Zvi; KANE, Alex; MARCUS, Alan J. Investments. 10. ed. New York: McGraw-Hill Education, 2014.
- MILTON, M.; ARTHUR, E. Practical Cluster Analysis for Data Scientists. Springer, 2017.
- FERREIRA, M. fundamentus: Python Fundamentus. 2021. Disponível em: <https://pypi.org/project/fundamentus/>, Abril 2025.
- FUNDAMENTUS. Ferramenta de análise fundamentalista de ações. Disponível em: <https://www.fundamentus.com.br/>, Abril 2025.
- PORTAL DATA SCIENCE. “Introdução à clusterização e os diferentes métodos”, 2023. Disponível em: <https://portaldatascience.com/introducao-a-clusterizacao-e-os-diferentes-metodos/>, Abril 2025.
- SUNO ANALÍTICA. Visualização de Ação Ambev S.A. (ABEV3). Disponível em: <https://www.suno.com.br/analitica/acoes/abev3/>, Abril 2025.
- STATUS INVEST. “Busca Avançada de Ações”, site Status Invest. Disponível em: <https://statusinvest.com.br/acoes/busca-avancada>, Abril 2025.
- INVESTIDOR 10. “Busca Avançada de Ações”, site Investidor 10. Disponível em: <https://investidor10.com.br/acoes/busca-avancada/>, Abril 2025.
- TORO INVESTIMENTOS. “Empresas listadas na bolsa B3”. Disponível em: <https://blog.toroinvestimentos.com.br/bolsa/empresas-listadas-bolsa-b3/>, Abril 2025.