

ANÁLISE COMPARATIVA DE ALGORITMOS DE APRENDIZADO DE MÁQUINA E REDES NEURAIS NA PREVISÃO DO MERCADO DE AÇÕES BRASILEIRO

Gabriel V. de CARVALHO¹; Wedson G. S. JUNIOR²

RESUMO

Este trabalho realiza uma análise comparativa entre algoritmos de Aprendizado de Máquina e Redes Neurais na previsão de preços do mercado de ações brasileiro. Utilizando dados históricos dos ativos CMIG4 e ITUB4 extraídos do Yahoo Finance, foram aplicados os modelos Random Forest Regressor, e LSTM. Os dados foram normalizados e tratados com ferramentas da linguagem Python, em ambiente Google Colab. O desempenho dos modelos foi avaliado com base em métricas e visualizações gráficas, permitindo identificar as abordagens mais eficazes. Os resultados indicam que as Redes Neurais, especialmente o modelo LSTM, apresentaram maior capacidade de prever padrões temporais nos dados analisados.

Palavras-chave: Redes Neurais; Aprendizado de Máquina; Mercado Financeiro; Previsão do Mercado Financeiro; Algoritmos de Previsão de Dados.

1. INTRODUÇÃO

De acordo com Silva (2024) o uso de Inteligência Artificial no setor financeiro tem crescido com o avanço de técnicas voltadas à previsão e análise de dados, mas ainda não oferece orientações totalmente confiáveis aos investidores. A alta volatilidade do mercado pode levar a decisões precipitadas, especialmente entre iniciantes, resultando em perdas significativas. Uma ferramenta de auxílio para a tomada de decisões, baseado em aprendizado de máquina, pode representar um passo importante para os investidores. O Aprendizado de Máquina tem ganhado destaque nas finanças pela capacidade de lidar com grandes volumes de dados e apoiar previsões de cenários financeiros (Nazareth; Reddy, 2023). Algoritmos como o Decision Tree Regressor e Random Forest Regressor, que define regras com base em covariáveis, são amplamente utilizados nesse contexto (Venkatasubramiam et al., 2017). As Redes Neurais Artificiais (RNAs), segundo Fazeli e Houghten (2019), apresentam grande potencial na previsão de valores e tendências de mercado. Dentre elas, a arquitetura LSTM se destaca por sua capacidade de capturar dependências de longo prazo em séries temporais, utilizando portões que controlam o fluxo de informações, o que a torna especialmente adequada para análise de dados históricos no mercado de ações (Vargas et al., 2022).

Por fim, este trabalho busca explorar os algoritmos citados e realizar uma comparação entre as diferentes arquiteturas, a fim de identificar aquela que apresenta o melhor desempenho na previsão dos valores, com base nos resultados obtidos por meio de treinamentos, testes e análises.

¹ Discente de Bacharelado em Ciência da Computação, IFSULDEMINAS – Campus Passos. E-mail: gabrielvenancio1123@gmail.com

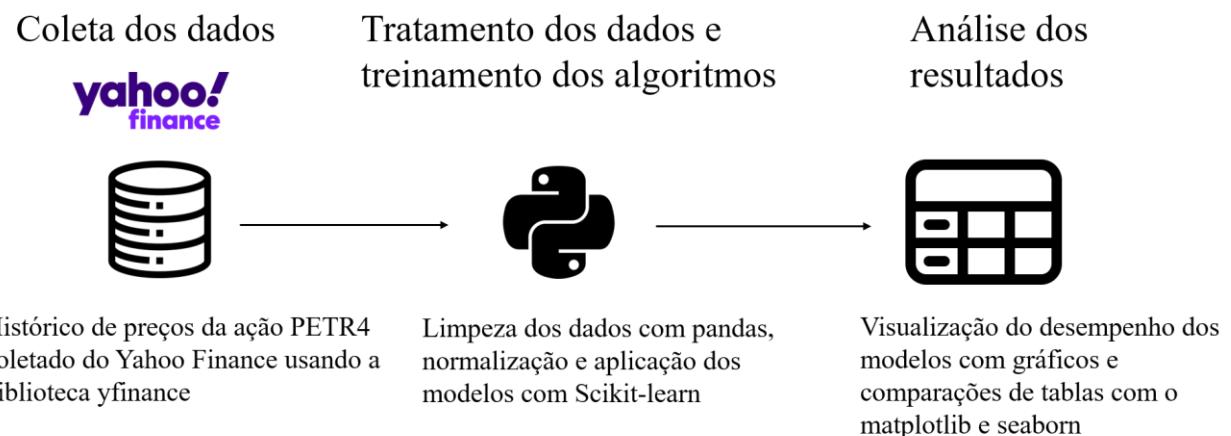
² Orientador, IFSULDEMINAS – Campus Passos. E-mail: wedson.junior@ifsuldeminas.edu.br

2. MATERIAL E MÉTODOS

Neste trabalho serão utilizados dados do Yahoo Finance, que oferece informações históricas sobre ativos do mercado financeiro. Os ativos analisados são ITUB4 (setor bancário) e CMIG3 (setor de energia), escolhidos sem critério específico, visando ampliar a análise por serem de setores distintos. As implementações serão feitas no Google Colab, ambiente em nuvem com suporte ao desenvolvimento em Python, ideal para aplicações de ciência de dados e aprendizado de máquina.

Com o uso da linguagem Python e bibliotecas como o yfinance, é possível obter dados históricos de ativos e organizá-los em dataframes, facilitando o acesso a informações como data, preço de abertura, fechamento, máximas, mínimas e volume. Apesar da variedade de variáveis disponíveis, a análise de previsão focará principalmente na variável Close, que representa o preço de fechamento diário. A Figura 1 ilustra o pipeline das etapas envolvidas nesse processo.

Figura 1 – Pipeline das atividades a serem realizadas



Fonte: Autoria Própria (2025).

O gerenciamento e a manipulação desses dataframes serão realizados por meio da biblioteca Pandas do Python, que permite trabalhar com dados tabulares de forma prática e eficiente. Para Aprendizado de Máquina, será usado o algoritmo Random Forest Regressor, por meio da biblioteca Scikit-learn do Python. Já para Redes Neurais, será usado o algoritmo Long Short-Term Memory, utilizando a biblioteca TensorFlow, também do Python.

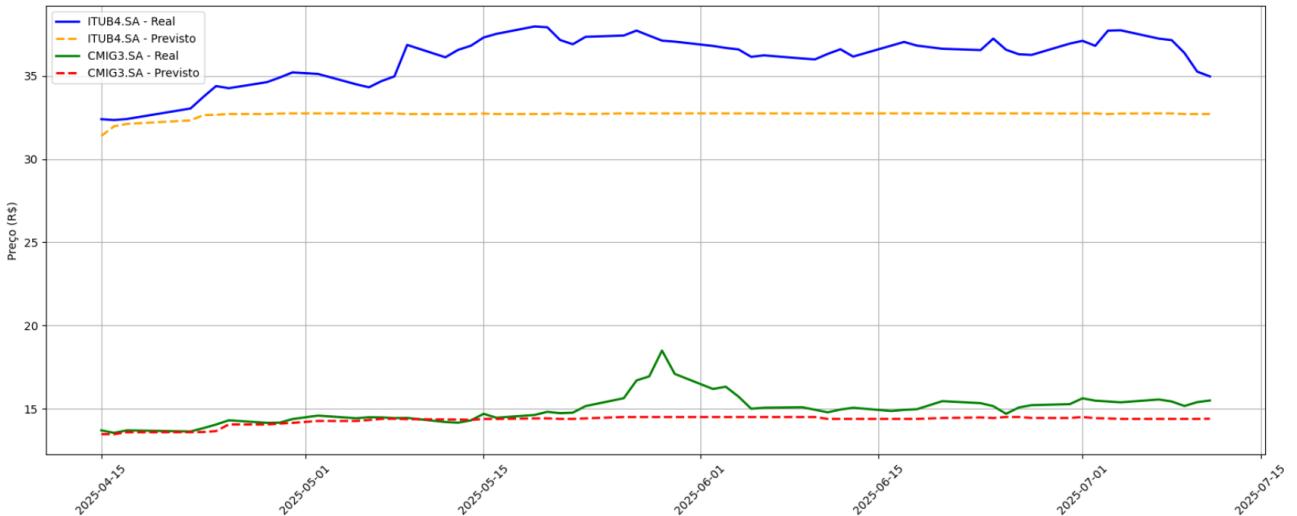
3. RESULTADOS E DISCUSSÃO

O modelo Random Forest apresentou boa assertividade para o ativo CMIG3 entre 15/04/2025 e 15/05/2025, mas teve desempenho ruim nos demais períodos, especialmente no fim de maio, quando não conseguiu prever uma alta repentina nos preços, pois fatores como mudança na taxa SELIC, divulgação de resultados da empresa ou até mesmo interferência política podem ter atrapalhado. As métricas indicam um Erro Absoluto Médio de R\$0,70, Erro Quadrático Médio de 1,01 e R² de -0,3180, mostrando que o modelo não explicou bem a variabilidade dos dados.

Para o ativo ITUB4, o desempenho foi ainda pior, sem acompanhar a tendência do mercado. As métricas foram: Erro Médio Absoluto de R\$3,44, Erro Quadrático Médio de 13,44 e R² de -5,8738,

evidenciando que o modelo falhou completamente. Com base nisso, conclui-se que o Random Forest não é adequado para prever o comportamento desses ativos. Os resultados podem ser analisados a baixo na Figura 2.

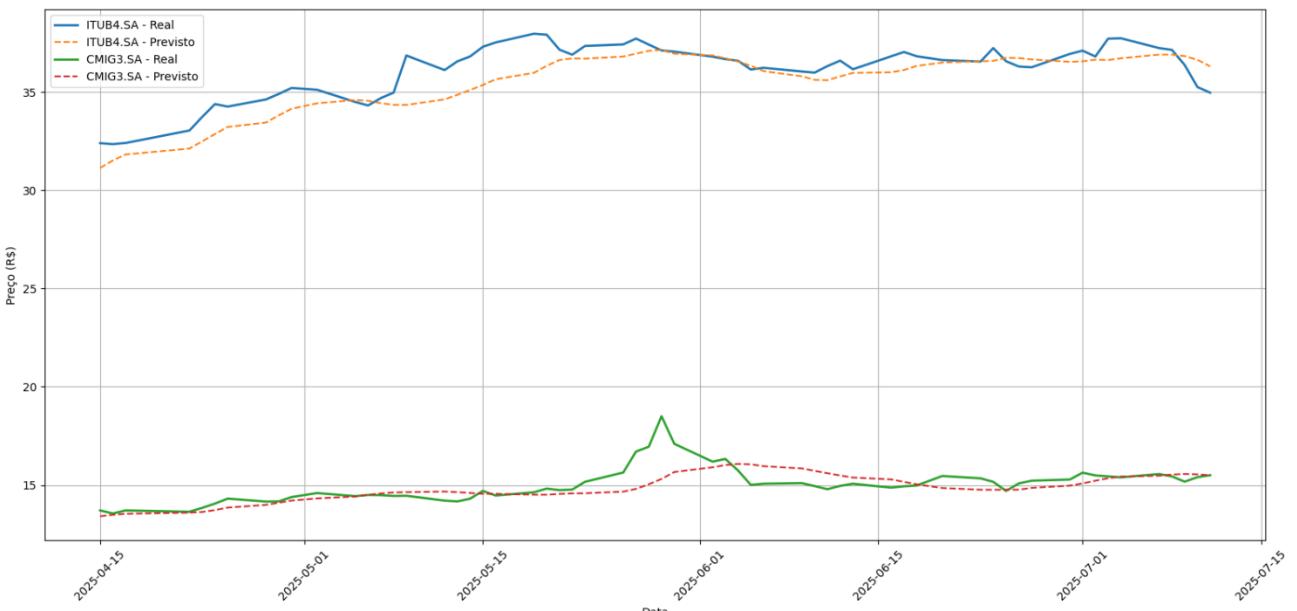
Figura 2- Previsão de Preço – Random Forest



Fonte: Autoria Própria (2025).

O modelo LSTM apresentou desempenho superior ao modelo anterior, conseguindo seguir melhor a tendência dos preços, mesmo sem prever com precisão a alta repentina no fim de maio. Para o ativo CMIG3, o Erro Médio Absoluto foi de R\$0,80, com Erro Quadrático Médio de 1,03 e R² de -0,3528, indicando dificuldades em explicar a variabilidade dos preços. Já para o ativo ITUB4, os resultados foram melhores: Erro Médio Absoluto de R\$0,98, Erro Quadrático Médio de 1,41 e R² de 0,27, mostrando que o modelo conseguiu captar parte da variação dos preços. Assim, a LSTM mostrou-se mais promissora para esse tipo de previsão. Os resultados podem ser observados na Figura 3.

Figura 3 – Previsão de Preço com LSTM



Fonte: Autoria Própria (2025).

4. CONCLUSÃO

A partir da análise comparativa entre algoritmos de Aprendizado de Máquina e Redes Neurais aplicados à previsão do mercado de ações brasileiro, foi possível observar que o algoritmo de Rede Neural, LSTM, conseguiu se sair melhor que o de aprendizado de máquina, uma vez que ele acompanha melhor séries temporais. Porém, ainda assim, não foi possível acompanhar de maneira 100% precisa o valor real dos ativos. Isso ocorre devido a diversos fatores externos ao preço histórico, como eventos políticos, econômicos, variações globais e até comportamento de investidores, que comprometem a previsão.

REFERÊNCIAS

Fazeli, A.; Houghten, S. Deep learning for the prediction of stock market trends. In: IEEE International Conference on Big Data, 2019, Los Angeles, p. 5513-5521.

Nazareth, N.; Reddy, Y. V. R. Financial applications of machine learning: A literature review. **Expert System with Applications**, v. 219, n. 119640, Junho 2023.

Tratkowski, Grzegorz. Identification of nonlinear determinants of stock indices derived by Random Forest algorithm. **International Journal of Management and Economics**, v. 56, n. 3, p. 209-217, Setembro 2020.

Vargas, Gabriel M. et al. B3 Stock Price Prediction Using LSTM Neural Networks and Sentiment Analysis. **IEEE Latin America Transactions**, v. 20, n. 7, p. 1067-1074, Julho 2022.

Venkatasubramaniam, Ashwini et al. Decision trees in epidemiological research. **Emerging Themes in Epidemiology**, v. 14, n. 11, Setembro 2017.