

ANÁLISE DO COMPORTAMENTO DE AÇÕES NA BOLSA DE VALORES UTILIZANDO APRENDIZADO DE MÁQUINAS

Lucas Breda Lima Mascarenhas ¹; Artur Henrique Kronbauer ².

¹ Graduando em Engenharia de Computação; Iniciação Científica - CNPq;
lucas.mascarenhas@aln.senaicimatec.edu.br.

² Centro Universitário SENAI CIMATEC; Salvador - BA; artur.kronbauer@fieb.org.br

RESUMO

A pesquisa proposta tem como justificativa investigar as variações e comportamentos de ações na bolsa de valores B3 (Bolsa, Brasil, Balcão) por meio da aplicação de técnicas de aprendizado de máquina (Machine Learning). Utilizando dados de volume de ações negociadas de uma empresa, valores de abertura e fechamento diários, a fim de entender o comportamento das ações. O modelo está sendo treinado com uma dezena de ações a fim de generalizar a rede neural e poder efetuar análises globais da B3.

PALAVRAS-CHAVE: Inteligência Artificial; Aprendizado de Máquinas, Predição de Ações; Bolsa de Valores.

1. INTRODUÇÃO

O projeto proposto tem como objetivo estudar as flutuações no comportamento do mercado financeiro brasileiro através da aplicação de técnicas de aprendizado de máquinas. O atual ambiente de mercado é caracterizado pela complexidade e volatilidade, o que exige abordagens inovadoras à análise e previsão de ativos financeiros. Os algoritmos de aprendizado de máquinas, mostram-se ferramentas poderosas de predição e análise de séries temporais financeiras. Ao analisar tendências históricas de preços, volumes de negociação, indicadores técnicos e eventos macroeconômicos, estas redes podem identificar relações complexas que não são possíveis utilizando métodos tradicionais¹. Esse recurso permite uma previsão mais precisa das tendências do mercado e contribui para a tomada de decisões estratégicas de investidores e empresas.

2. METODOLOGIA

O caráter aplicado desta pesquisa visa gerar conhecimento prático para uma análise eficaz do comportamento dos ativos no mercado financeiro. Aplicando uma abordagem quantitativa, a investigação centra-se na recolha e análise de dados numéricos e mensuráveis, utilizando técnicas estatísticas para estabelecer relações de causa, efeito e generalizações entre variáveis.

Para a realização desta pesquisa foi selecionado, inicialmente, o ativo "PETR4" (referente à Petrobrás) devido à sua alta liquidez no mercado financeiro brasileiro². Foram coletados dados de cotação do ativo a partir de 24 de março de 2006 até 17 de abril de 2024, com auxílio da API YFinance (Interface de Programação de Aplicação do Yahoo Finanças), totalizando 6682 análises diárias contendo informações de todos os indicadores escolhidos.

Para além das cotações da ação da Petrobrás, foram carregados, no mesmo intervalo de tempo, o volume de transações e o valor de abertura do ativo, o índice Ibovespa e a cotação do dólar americano. Após isso, a base de dados foi tratada utilizando a técnica "Imputação de Dados Faltantes" de maneira a excluir linhas com alguma informação inexistente (como em casos de feriados ou finais de semana quando a bolsa de valores não opera), para que os algoritmos apresentem resultados mais assertivos, deixando-a com 4436 linhas de dados.

Por conseguinte, 75% (3327 linhas) da base de dados foi designada para o treinamento dos algoritmos e 25% (1109 linhas) foi designada para o teste dos algoritmos. Também foram estudadas e avaliadas as precisões de 5 modelos de aprendizado de máquina de regressão distintos, a fim de estabelecer qual deveria ser utilizado para o desenvolvimento das próximas etapas da pesquisa. Foram ponderados os

métodos de regressão múltipla, polinomial, por árvore de decisão, por “random forest” e por redes neurais. Cada sistema foi testado e avaliado com base na taxa de sucesso na previsão do valor da cotação da ação do dia subsequente e seu erro absoluto médio (MAE).

Tabela 1: Algoritmos X Precisão Média X MAE Médio

Algoritmo	Precisão Média (%)	MAE Médio (±) em R\$
Regressão Múltipla	99,84491	0,17561
Regressão Polinomial	99,84285	0,17588
Regressão por Árvore de Decisão	99,62808	0,06598
Regressão por Random Forest	99,79544	0,10473
Regressão por Redes Neurais	99,81417	0,27487

Fonte: O autor

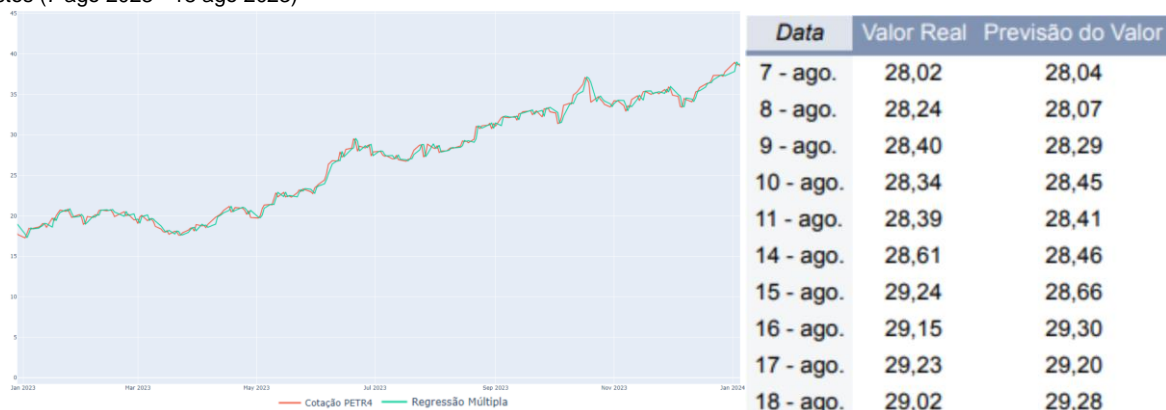
Observa-se que o método de Regressão Múltipla apresenta uma maior precisão média, com 99.84491% de precisão, quando comparado com os outros algoritmos, portanto, será selecionado para garantir melhor acurácia nos resultados almejados.

3. RESULTADOS E DISCUSSÃO

A análise dos movimentos do mercado de ações é uma área importante no cenário financeiro, onde a capacidade de prever corretamente os movimentos do mercado pode gerar lucros significativos ou minimizar perdas significativas. Tradicionalmente, esta análise utiliza métodos estatísticos e análises técnicas. No entanto, devido à crescente complexidade e volatilidade dos mercados financeiros, tornou-se clara a necessidade de abordagens mais avançadas³. Neste cenário, os algoritmos de aprendizado de máquinas aparecem como uma ferramenta promissora para melhorar a precisão da previsão da volatilidade das ações.

A Figura 1 e a Tabela 2 demonstram os resultados da predição de valores após o treinamento inicial do algoritmo de regressão múltipla comparado com os valores reais da cotação da Petrobrás (PETR4) no intervalo de tempo de um ano (Figura 1) e 10 dias (Tabela 2).

Figura 1: Cotação PETR4 X Previsão por Regressão Múltipla (jan 2023 - jan 2024); Tabela 2: Comparativo entre valores reais e valores previstos (7 ago 2023 - 18 ago 2023)



Fonte: O autor

Fonte: O autor

O aprendizado de máquina se destaca por sua capacidade de analisar grandes conjuntos de dados e identificar padrões complexos, que os métodos tradicionais podem não deduzir⁴. Por meio de algoritmos como o de regressão múltipla, o aprendizado de máquina fornece maior destreza ao analisar o mercado de ações, permitindo previsões mais precisas que podem se adaptar à dinâmica real através da capacidade de generalização do algoritmo.

Utilizando métodos quantitativos, dados numéricos são coletados e analisados, aplicando técnicas estatísticas para estabelecer relações de causa e efeito entre variáveis, identificar padrões e tendências de mercado e fornecer previsões precisas dos movimentos das ações⁵.

4. CONSIDERAÇÕES FINAIS

A análise, interpretação e previsão do comportamento de ações/ativos na bolsa de valores brasileira, através de um algoritmo de aprendizado de máquinas, requer ajustes minuciosos a fim de alcançar resultados satisfatórios. Portanto, a escolha do algoritmo de regressão múltipla foi essencial para o desenvolvimento atual e futuro da pesquisa, uma vez que, atingiu-se resultados satisfatórios e significativos.

A escolha adequada das variáveis preditoras foi de forte importância para o sucesso dos resultados atuais⁶, visto que, o método de machine learning adotado e treinado teve acurácia acima de 99% nos testes assistidos.

A grande quantidade de informações, mesmo após o tratamento da base de dados, foi essencial para um bom treinamento do método proposto⁷, uma vez que o método foi capaz de acompanhar de forma regular o gráfico de valores reais, tanto em aumentos/quedas bruscas, quanto em momento lineares vistos na cotação do ativo.

Como futuros trabalhos, busca-se a generalização do algoritmo⁸ para utilização em outros ativos, a verificação do desempenho de outros algoritmos de machine learning e o desenvolvimento de uma aplicação que, dado os valores de investimento e o tempo de aplicação, retorne como resultado a estimativa de ganho/perda num determinado ativo.

5. REFERÊNCIAS

¹ FERREIRA, Mariana. **Os desafios da Machine Learning: Aplicação ao Mercado Financeiro**. Coimbra, 2018.

² PEREIRA, Eunice. **Aplicação de Redes Neurais Artificiais na Predição de Preços de Ações por Indicadores Financeiros**. Uberlândia, 2018.

³ GÓES, Guilherme. **Redes neurais artificiais na predição do preço de ações**. Florianópolis, 2019.

⁴ CARVALHO, Gustavo. **Algoritmos De Machine Learning Para Previsão De Ações Da B3**. Uberlândia, 2020.

⁵ GALVÃO, Guilherme. **Modelo de Predição de Preços de Ações Utilizando Redes Neurais Artificiais**. São Paulo, 2022.

⁶ SANTANA, Pedro. **A Influência de Variáveis Preditoras na Performance de Redes Neurais Recorrentes para Estimativa do Preço de Ações Integrantes do Ibovespa**. Brasília, 2023.

⁷ PRADO, Gustavo. **Pré-processamento de Dados em Aprendizado de Máquina Supervisionado**. São Carlos, 2003.

⁸ SAENGYONGAM, Sorawit. **Effect-Invariant Mechanisms for Policy Generalization**. Zurique: Journal of Machine Learning Research, 2024.