

Relatorio Tecnico - Projeto de Classificacao de Precos de Acoes

Autor: William Brandao

Profissao: Cientista de Dados

Data: 04/06/2025

1. Introducao

O presente relatorio tecnico descreve o desenvolvimento de um projeto de modelagem preditiva voltado a classificacao de acoes com base em dados contabeis. O objetivo central e prever tendencias de precos de acoes utilizando variaveis fundamentalistas, com foco na construcao de modelos capazes de gerar vantagem estatistica para alocao de capital, mesmo sem a necessidade de prever o preco exato dos ativos. A abordagem visa apoiar decisoes quantitativas e fundamentadas para investidores institucionais e gestores de portfolio.

2. Metodologia

Foram utilizadas variaveis contabeis trimestrais, incluindo indicadores como Receita Liquida, EBITDA, Lucro Liquido, Ativos, Passivos, Patrimonio Liquido, entre outros. A lista completa de variaveis contempla mais de 25 metricas financeiras, detalhadas no anexo do projeto. Os targets considerados foram o preco de fechamento das acoes e uma classificacao binaria (valorizacao ou nao no periodo seguinte). As abordagens metodologicas incluíram validacao cruzada, controle de overfitting, ajuste de hiperparametros e normalizacao dos dados. Entre os principais desafios, destacam-se o baixo poder explicativo (R^2 proximo de zero) e metricas de erro elevadas na previsao direta de precos, reforcando a necessidade de abordagens baseadas em classificacao e ranking.

3. Transformacoes com Indicadores Financeiros

Para aprimorar a robustez e comparabilidade dos modelos, foram calculados indicadores derivados, tais como:

- Margem EBIT = $EBIT / Receita\ Liquida$
- ROE = $Lucro\ Liquido / Patrimonio\ Liquido$
- ROA = $Lucro\ Liquido / Ativo\ Total$
- Margem Liquida = $Lucro\ Liquido / Receita\ Liquida$
- Alavancagem Financeira = $Ativo\ Total / Patrimonio\ Liquido$
- Cobertura de Juros = $EBIT / Despesas\ Financeiras$

Essas transformacoes permitem comparar empresas de diferentes portes e setores, reduzem o ruido dos dados brutos e aumentam a robustez dos modelos preditivos.

4. Resultados

A performance dos modelos foi avaliada por meio da simulacao de duas carteiras: uma com pesos iguais (carteira igualitaria) e outra quantitativa, com alocao baseada nas probabilidades

estimadas pelo modelo de classificacao. Os resultados demonstram que a carteira quantitativa supera consistentemente a igualitaria ao longo de multiplos anos, evidenciando a eficacia da abordagem baseada em dados contabeis.

A seguir, apresenta-se o grafico comparativo das estrategias.

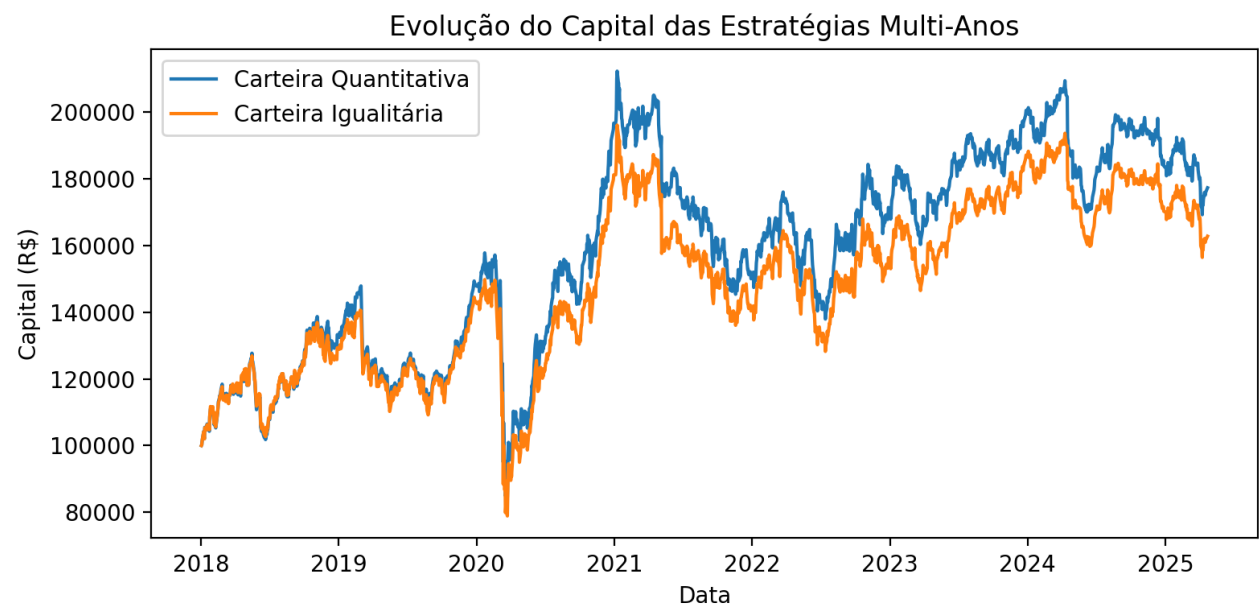


Figura 1 - Evolucao do capital de duas carteiras com 10 acoes: a Carteira Quantitativa (pesos alocados com base no modelo preditivo) supera consistentemente a Carteira Igualitaria (pesos fixos iguais) ao longo de multiplos anos, indicando a eficacia da abordagem baseada em dados contabeis.

5. Discussao: Pontos Positivos e Limitantes

Entre os pontos positivos do projeto, destacam-se a estrutura modular e automatizada do codigo, a clareza documental, a possibilidade de expansao para outros ativos e a entrega parametrizavel em formato PDF. O pipeline desenvolvido permite facil manutencao e reproduzibilidade dos resultados.

Como limitacoes, ressalta-se a dificuldade inerente de prever precos de acoes diretamente a partir de dados contabeis, dada a influencia de fatores externos e macroeconomicos. A integracao de variaveis externas, como indicadores macroeconomicos, setoriais e de sentimento de mercado, e recomendada para aprimorar a acuracia dos modelos.

6. Conclusao e Proximos Passos

O projeto demonstrou que, mesmo sem prever precos com alta precisao, e possivel construir estrategias quantitativas robustas e superiores a alocao tradicional de pesos iguais, utilizando apenas dados contabeis. A abordagem baseada em classificacao e ranking de empresas mostrou-se eficaz para geracao de alfa e otimizacao de portfolios.

Como próximos passos, recomenda-se a integração de dados macroeconômicos, setoriais e de sentimento, além do aprimoramento do tratamento de dados faltantes e outliers. A expansão do pipeline para outros mercados e a avaliação de diferentes algoritmos de machine learning também são sugeridas para futuras iterações.