

PREVISÃO DIÁRIA DO IBOVESPA COM REGRESSÃO LINEAR

TECH CHALLENGE FASE 2



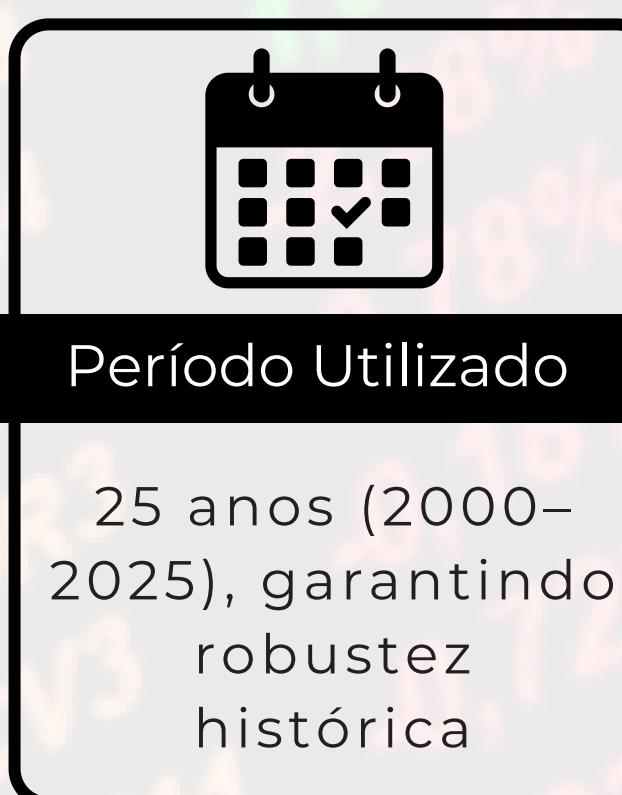
1^a ETAPA

AQUISIÇÃO E EXPLORAÇÃO DOS DADOS



Fonte

Fonte: Yahoo Finance (via biblioteca yfinance)



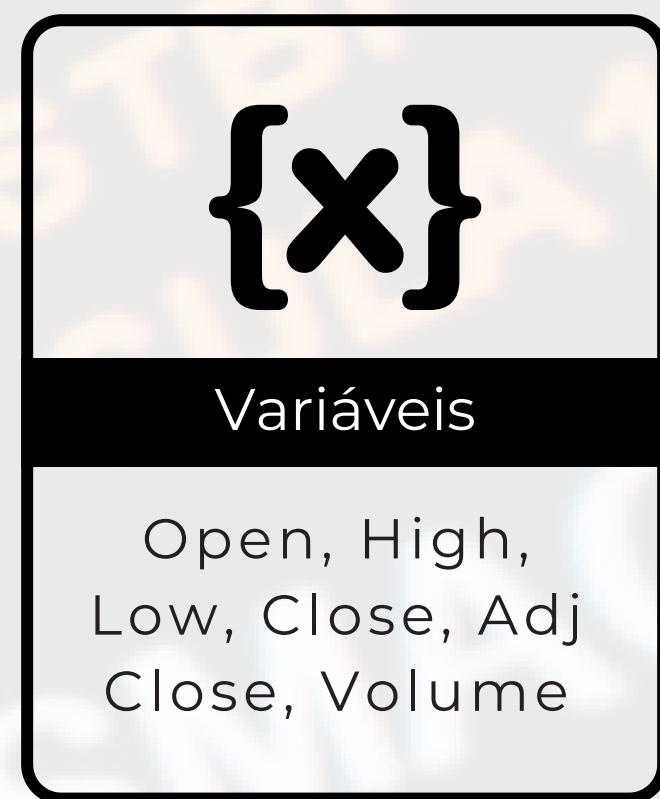
Período Utilizado

25 anos (2000–2025), garantindo robustez histórica



Frequência

Diária



Variáveis

Open, High, Low, Close, Adj Close, Volume

Durante a exploração inicial, observamos tendência de longo prazo e ciclos de volatilidade, indicando não estacionariedade — o que exigiu transformações específicas para estabilizar a série.

MODELAGEM E COMPARAÇÃO DE DESEMPENHO

Foram avaliadas três abordagens principais:

- **Regressão Logística** – baseline linear supervisionado, com `class_weight='balanced'`.
- **XGBoost** – ensemble de árvores de decisão com boosting gradiente, ajustado via `scale_pos_weight`.
- **LSTM** – exploratórios, para análise comparativa e validação estatística.

Modelo	Acurácia	Observações
Regressão Logística	76,66%	Atendeu ao requisito mínimo de acurácia e apresentou comportamento estável.
XGBoost	70%	Ficou abaixo do limite exigido
LSTM	—	$R^2 \approx 75\%$; erro médio (RMSE) ≈ 4.24 ; útil como validação estatística complementar.

MODELAGEM E COMPARAÇÃO DE DESEMPENHO

Foram avaliadas três abordagens principais:

- Durante a análise, ficou evidente que a maioria dos últimos 30 dias foi de quedas, o que resultou em forte desbalanceamento de classes.
- O XGBoost, mesmo com boa sensibilidade, não atingiu o limite mínimo de acurácia exigido.
- Já a **Ressão Logística**, embora simples, apresentou melhor estabilidade e superou o requisito, sendo o modelo final escolhido.

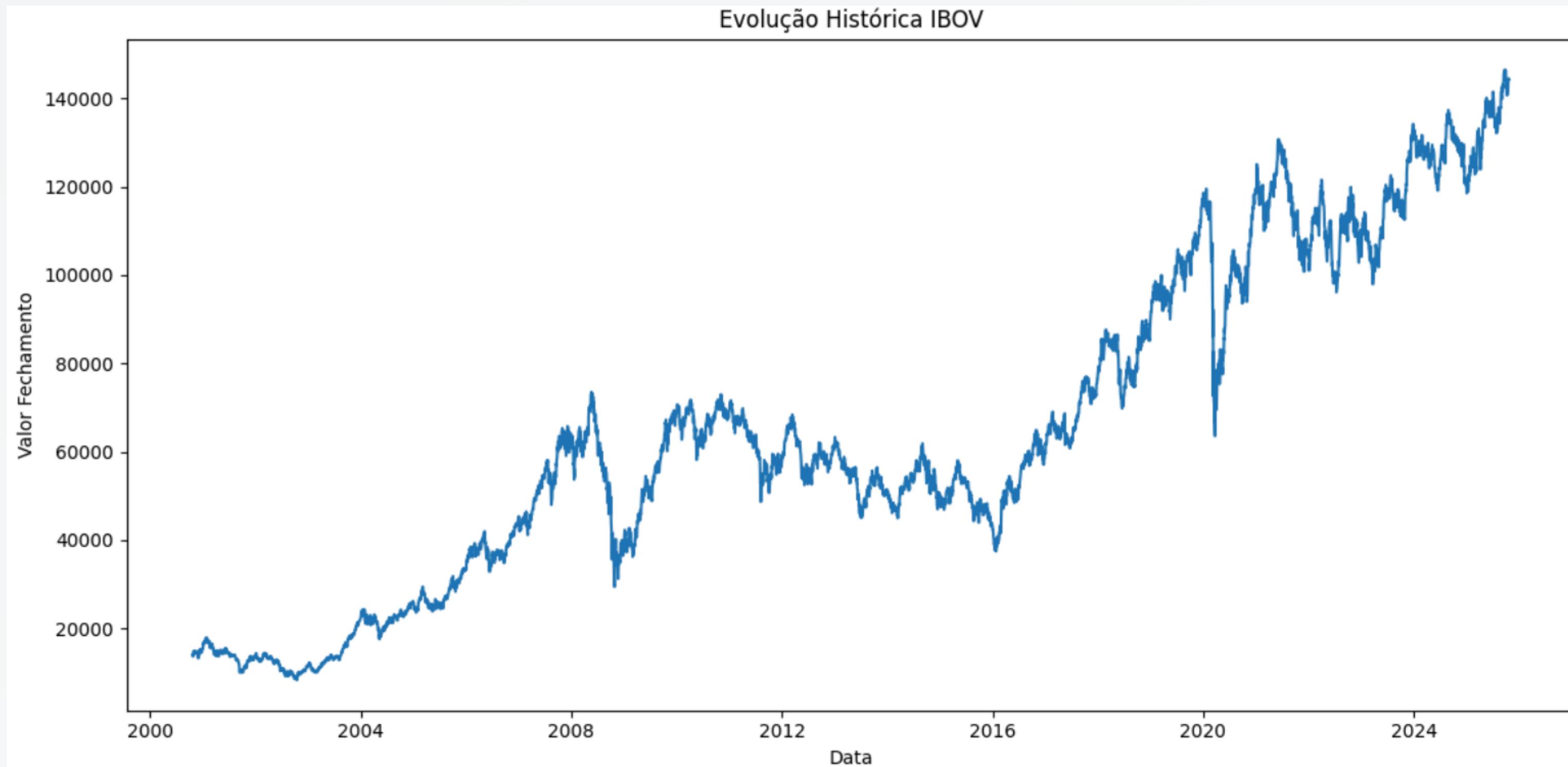
Modelo	Acurácia	Observações
Ressão Logística	76,66%	Atendeu ao requisito mínimo de acurácia e apresentou comportamento estável.
XGBoost	70%	Ficou abaixo do limite exigido
LSTM	—	$R^2 \approx 75\%$; erro médio (RMSE) ≈ 4.24 ; útil como validação estatística complementar.

RESULTADOS E AVALIAÇÃO DE MÉTRICAS

Modelo	Acurácia	Recall (Classe 1)	F1-score
Regressão Logística	76,66	1	85
XGBoost	70	62	42,85

- Acurácia em teste: 76,66% ✓
- Matriz de confusão (teste): [[20,3],[6,1]]

GRÁFICOS COMPLEMENTARES:

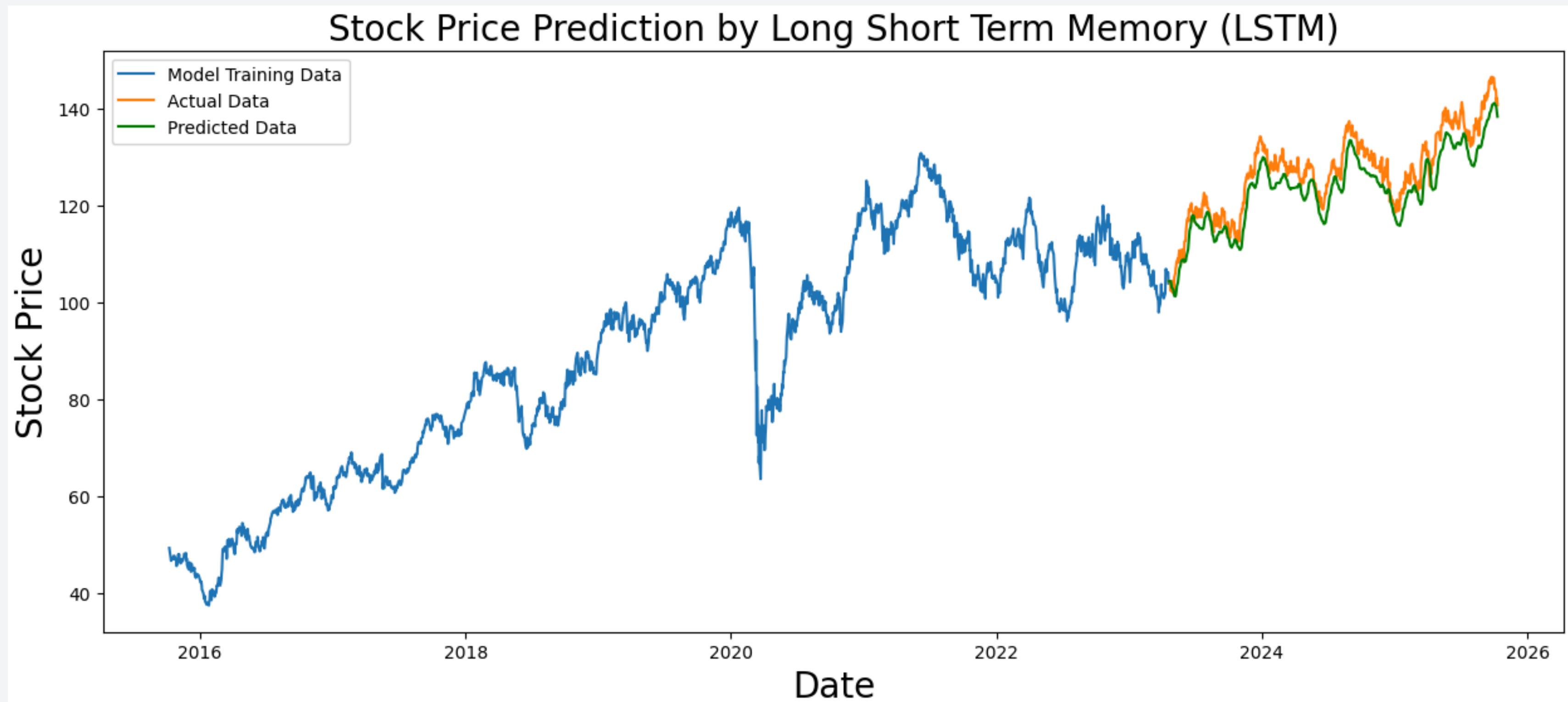


GRÁFICOS COMPLEMENTARES:

Modelos Xgboost e Regressão Logística



GRÁFICOS COMPLEMENTARES:



CONCLUSÕES E IMPACTO DE NEGÓCIO

- O modelo baseado na Regressão Logística, atingiu **76,66%** de acurácia no conjunto de teste, atendendo ao requisito mínimo estabelecido.
- O modelo se mostrou confiável para uso em dashboards internos e fornece insumo quantitativo para decisões diárias da mesa de investimentos.
- O trabalho também consolidou um pipeline completo e reproduzível, que permite futuras melhorias e integração com estratégias automatizadas.

OBRIGADO!