

# Desafio Técnico 2 - Direção do Ibovespa

Análise e Previsão da Direção do Ibovespa  
para o Próximo Dia

Cinthia Gonzalez da Silva  
Gabriel Huzian  
Karyne Barbosa Silva



## Modelos, Validação e Métricas

Análise detalhada dos modelos testados, validação e métricas de desempenho.



## Features e Preparação dos Dados

## Objetivo do Projeto e Regras Fundamentais



# Objetivo do Projeto e Regras Fundamentais

O nosso desafio consiste em prever a direção do Ibovespa para o dia seguinte, identificando se o índice fechará acima do dia atual (classe '1, sobe') ou não (classe '0, não sobe'). As regras são claras: utilizaremos apenas os dados disponíveis no CSV fornecido, separaremos um conjunto de teste com os últimos 30 pregões e estabelecemos uma meta de acurácia mínima de 75% para os modelos, priorizando o F1 em caso de empates.





## RETORNOS

Os RETORNOS calculados diariamente, em janelas de 5 e 10 dias, ajudam a identificar o momentum do mercado, indicando se a tendência é de alta ou baixa ao longo do tempo.

## MÉDIAS MÓVEIS

As MÉDIAS MÓVEIS de 5, 10 e 20 dias fornecem uma visão dos preços médios ao longo do tempo, permitindo identificar padrões de movimento e possíveis pontos de reversão.

## AFASTAMENTOS

Os AFASTAMENTOS do preço em relação às médias móveis ajudam a identificar se o preço está supervalorizado ou subvalorizado, sinalizando potenciais correções.

## VOLATILIDADE

A VOLATILIDADE, em janelas de 5 e 10 dias mede a variação dos preços, indicando períodos de maior incerteza e possíveis riscos no mercado.

## RSI (Índice de Força Relativa)

O RSI de 14 períodos mede a força e a velocidade do movimento de preços, ajudando a identificar condições de sobrecompra ou sobrevenda no mercado.

## Stochastic %K

O Stochastic %K (14,3) indica a posição do preço atual em relação à sua faixa de preços recente, ajudando a identificar potenciais reversões de tendência.

## VARIAÇÃO DO VOLUME

A variação do VOLUME dia a dia é crucial para entender a força de uma tendência, pois um aumento no volume pode confirmar movimentos de preços significativos.

## VARIÁVEIS DE CALENDÁRIO

As variáveis de calendário, como dia da semana, permitem capturar padrões sazonais que podem impactar o comportamento do mercado.



# Features e Preparação dos Dados



Prezi

# RETORNOS

Os RETORNOS calculados diariamente, em janelas de 5 e 10 dias, ajudam a identificar o momentum do mercado, indicando se a tendência é de alta ou baixa ao longo do tempo.



# MÉDIAS MÓVEIS

As MÉDIAS MÓVEIS de 5, 10 e 20 dias fornecem uma visão dos preços médios ao longo do tempo, permitindo identificar padrões de movimento e possíveis pontos de reversão.

# AFASTAMENTOS

Os AFASTAMENTOS do preço em relação às médias móveis ajudam a identificar se o preço está supervalorizado ou subvalorizado, sinalizando potenciais correções.



# **VOLATILIDADE**

A VOLATILIDADE em janelas de 5 e 10 dias mede a variação dos preços, indicando períodos de maior incerteza e possíveis riscos no mercado.



# RSI (Índice de Força Relativa)

O RSI de 14 períodos mede a força e a velocidade do movimento de preços, ajudando a identificar condições de sobrecompra ou sobrevenda no mercado.





# Stochastic %K

O Stochastic %K (14,3) indica a posição do preço atual em relação à sua faixa de preços recente, ajudando a identificar potenciais reversões de tendência.

# VARIAÇÃO DO VOLUME

A variação do VOLUME dia a dia é crucial para entender a força de uma tendência, pois um aumento no volume pode confirmar movimentos de preços significativos.

# VARIÁVEIS DE CALENDÁRIO

As variáveis de calendário, como dia da semana, permitem capturar padrões sazonais que podem impactar o comportamento do mercado.



# Modelos, Validação e Métricas

Análise detalhada dos modelos testados, validação e métricas de desempenho.

## Modelo 1: Regressão Logística

Testamos a Regressão Logística para avaliar relações lineares entre as variáveis e a direção do furosp.

## Modelo 2: SVM Linear

Implementamos o SVM Linear, que se destaca em classificações lineares e simplistas.

## Modelo 3: SVM com kernel RBF

O SVM com kernel RBF foi testado para capturar relações não lineares e complexas entre as features.

## Modelo 4: KNN

O KNN foi aplicado para entender a proximidade dos dados e suas classificações baseadas em vizinhos.

## Modelo 5: Árvore de Decisão

Utilizamos a Árvore de Decisão para visualizar decisões e suas consequências de forma hierárquica.

## Modelo 6: Random Forest

O Random Forest foi testado para melhorar a robustez e reduzir overfitting por meio de múltiplos árvores.

## Validação: TimeSeriesSplit

Aplicamos TimeSeriesSplit para validar os modelos, garantindo a integridade temporal dos dados durante o treinamento.

## Métricas-chave

As métricas utilizadas foram: Acurácia, Precisão, Recall, F1 e AUC, essenciais para avaliar o desempenho dos modelos.





# Modelo 1: Regressão Logística

Testamos a Regressão Logística para avaliar relações lineares entre as variáveis e a direção do Ibovespa.

## Modelo 2: SVM Linear

Implementamos o SVM Linear, que se destaca em classificações lineares e simplistas.

## Modelo 3: SVM com kernel RBF

O SVM com kernel RBF foi testado para capturar relações não lineares e complexas entre as features.

# Modelo 4: KNN

O KNN foi aplicado para entender a proximidade dos dados e suas classificações baseadas em vizinhos.



# Modelo 5: Árvore de Decisão

Utilizamos a Árvore de  
Decisão para visualizar  
decisões e suas  
consequências de forma  
hierárquica.

# Modelo 6: Random Forest

O Random Forest foi testado para melhorar a robustez e reduzir overfitting por meio de múltiplas árvores.

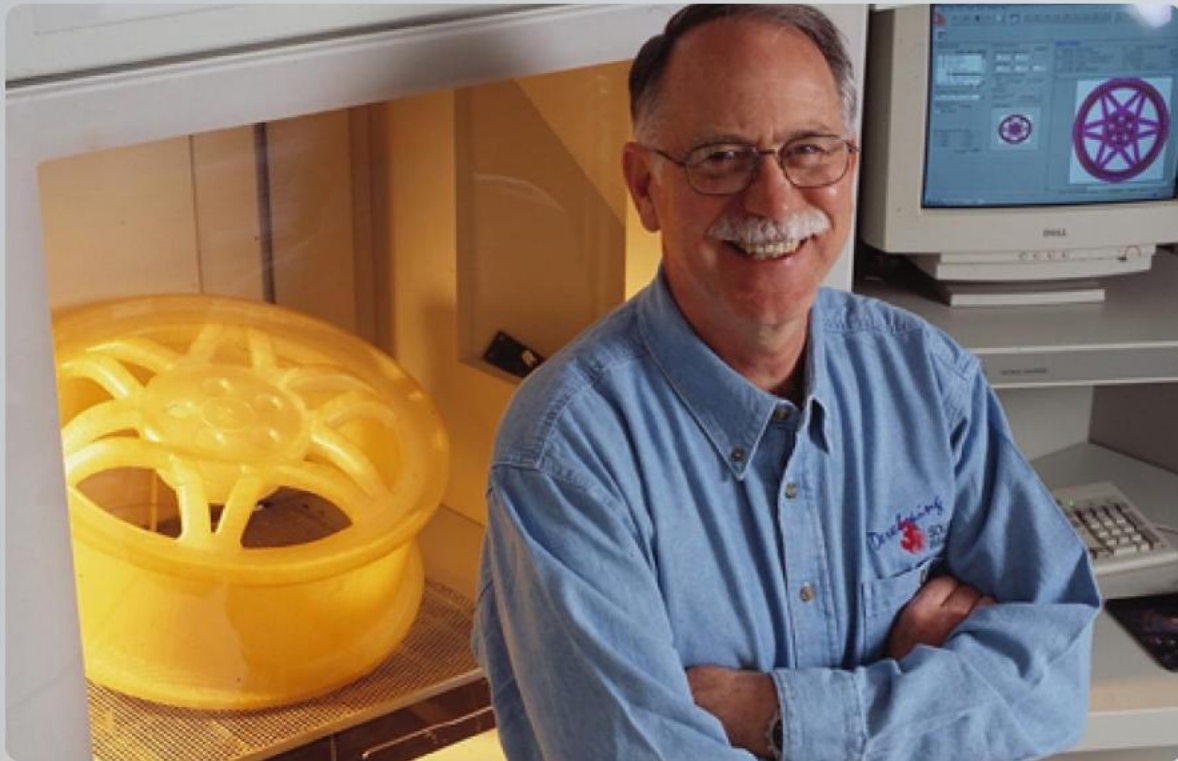
# Validação: TimeSeriesSplit

Aplicamos TimeSeriesSplit para validar os modelos, garantindo a integridade temporal dos dados durante o treinamento.

# Métricas-chave

As métricas utilizadas foram  
Acurácia, Precisão,  
Revocação, F1 e AUC,  
essenciais para avaliar o  
desempenho dos modelos.





## Modelo SVM com Kernel RBF: Resultados

O modelo SVM com kernel RBF se destacou, alcançando uma acurácia de aproximadamente 80%. A previsão para o dia 22/10/2025 indica que a probabilidade do Ibovespa subir no dia seguinte é inferior a 50%, sugerindo uma expectativa de estabilidade ou queda.



## Próximos Passos para Melhoria

Os próximos passos incluem a incorporação de custos de transação e slippage nas simulações, a testagem de diferentes limiares de decisão para melhorar a performance do modelo, a exploração de novos regimes de mercado como variáveis adicionais, e a realização de um backtest walk-forward para avaliar a robustez das previsões em múltiplas janelas de teste.