

ANA RAQUEL

POSTECH

DATA ANALYTICS

TECH CHALLENGE

FASE 02

TECH CHALLENGE

Tech Challenge é o projeto que englobará os conhecimentos obtidos em todas as disciplinas da fase. Esta é uma atividade que, em princípio, deve ser desenvolvida em grupo. Importante atentar-se ao prazo de entrega, pois trata-se de uma atividade obrigatória, uma vez que sua pontuação se refere a **90%** da nota final.

O problema

Você foi recentemente alocado em uma equipe de cientistas de dados de um grande fundo de investimentos brasileiro. Sua missão inicial é desenvolver um modelo preditivo capaz de prever **se o índice IBOVESPA vai fechar em alta ou baixa no dia seguinte**, com base em dados históricos do próprio índice. Esse modelo será usado como insumo para alimentar dashboards internos de tomada de decisão dos analistas quantitativos da empresa.

Dados

Utilize os dados históricos do índice IBOVESPA, disponíveis publicamente: <https://br.investing.com/indices/bovespa-historical-data>

Requisitos:

- Selecione o **período “diário”**.
- Baixe um intervalo de pelo menos **2 anos** de dados.
- Realize o pré-processamento necessário para utilizar os dados no modelo.

Objetivo

Criar um modelo que **preveja se o fechamento do IBOVESPA do dia seguinte será maior ou menor que o do dia atual**, ou seja:

Dia	Fechamento	Tendência
D1	120.000	—
D2	120.500	↑
D3	119.900	↓

Seu modelo deve prever a **tendência** (↑ ou ↓) com **acuracidade mínima de 75%** em um conjunto de teste. O conjunto de testes deverá conter o último mês (30 dias) de dados disponíveis.

Entregas Esperadas

Você deverá apresentar ao time de investimentos:

Storytelling Técnico

Apresentação com os seguintes requisitos:

- Aquisição e exploração dos dados;
- Estratégia de engenharia de atributos, caso seja realizado (por exemplo: variações, médias móveis, etc)
- Preparação da base para previsão (definição de target, janela de tempo, etc.);
- Escolha e justificativa do modelo utilizado;
- Resultados e análise de métricas que garantem que o modelo está confiável;

Justificativa Técnica

Explique claramente:

- Por que você escolheu o(s) modelo(s) utilizados (ex: árvores, regressão logística, LSTM, etc.);
- Como tratou a natureza sequencial dos dados (ex: janelas deslizantes, features lagged);
- Quais os trade-offs entre acuracidade e overfitting.

Quais documentos devo subir na plataforma? As entregas deverão conter necessariamente:

- Uma pasta zip contendo todos os arquivos do projeto, ou um link para o repositório do GitHub com seu projeto.
- Apresentação do storytelling. Pode ser em PPT, PDF ou ferramentas da sua preferência, como o CANVAS, por exemplo.
- Vídeo de no máximo **5 (cinco)** minutos, com uma visão gerencial, explicando como interpretar os resultados obtidos. Note que é necessário que todos os membros do grupo apareçam no vídeo.

Dica: Utilize o maior intervalo de tempo possível para atingir maior acuracidade no modelo preditivo.

Lembre-se que você poderá apresentar o desenvolvimento do seu projeto durante as lives e grupos de estudo com os(as) docentes. Essas são ótimas oportunidades para discutir sobre as dificuldades encontradas, além de pegar dicas valiosas com especialistas e colegas de turma.

Não se esqueça que isso é um entregável obrigatório! Se atente para o prazo de entrega até o final da fase.

Boa sorte!

The background is a dark blue field filled with numerous small, light blue dots, resembling a starry sky. Overlaid on this are several large, wavy, translucent bands in shades of teal, blue, and yellow. These bands flow from the left side towards the right, creating a sense of motion. Scattered throughout the composition are various geometric elements: a thin vertical line on the left, a circle containing the number '7' in the upper center, a circle with an 'X' inside on the lower left, a small yellow circle with an 'O' inside near the bottom center, and a yellow hexagon in the bottom right corner. The overall aesthetic is futuristic and technological.

POSTECH