

TECH CHALLENGE

Olá estudante! 😊

Tech Challenge é o projeto que engloba 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: Imagina que você foi contratado como cientista de dados de um hospital e tem o desafio de **desenvolver um modelo de Machine Learning para auxiliar os médicos e médicas a prever se uma pessoa pode ter obesidade.**

A obesidade é uma condição médica caracterizada pelo acúmulo excessivo de gordura corporal, a ponto de prejudicar a saúde. Esse problema tem se tornado cada vez mais prevalente em todo o mundo, atingindo pessoas de todas as idades e classes sociais. As causas da obesidade são multifatoriais e envolvem uma combinação de fatores genéticos, ambientais e comportamentais.

Utilizando a base de dados disponibilizada neste desafio em **“obesity.csv”**, desenvolva um modelo preditivo e crie um sistema preditivo para auxiliar a tomada de decisão da equipe médica a diagnosticar a obesidade.

Dicionário de dados:

- Gender: Gênero.
- Age: Idade.
- Height: Altura em metros.
- Weight: Peso em kgs.
- family_history: Algum membro da família sofreu ou sofre de excesso de peso?
- FAVC: Você come alimentos altamente calóricos com frequência?
- FCVC: Você costuma comer vegetais nas suas refeições?

- NCP: Quantas refeições principais você faz diariamente?
- CAEC: Você come alguma coisa entre as refeições?
- SMOKE: Você fuma?
- CH2O: Quanta água você bebe diariamente?
- SCC: Você monitora as calorias que ingere diariamente?
- FAF: Com que frequência você pratica atividade física?
- TER: Quanto tempo você usa dispositivos tecnológicos como celular, videogame, televisão, computador e outros?
- CALC: Com que frequência você bebe álcool?
- MTRANS: Qual meio de transporte você costuma usar?
- Obesity_level (coluna alvo): Nível de obesidade

Como entregável, você precisar desenvolver toda a pipeline de Machine Learning, garantindo a entrega de todos os seguintes requisitos listados a seguir que serão avaliados na sua entrega:

- ✓ **Pipeline de machine learning** demonstrando toda a etapa de feature engineering e treinamento do modelo.
- ✓ Modelo com **assertividade acima de 75%**.
- ✓ Realizar o **deploy do seu modelo em uma aplicação preditiva utilizando o Streamlit**.
- ✓ Construir uma **visão analítica em um painel com principais insights obtidos com o estudo** sobre obesidade para trazer insights para a equipe médica.
- ✓ **Compartilhar o link da sua aplicação deployada no app do Streamlit + link do painel analítico + link do repositório do seu github com todo código desenvolvido em um arquivo .doc ou .txt para realizar o upload na plataforma online.**
- ✓ Gravar um vídeo mostrando toda a estratégia utilizada e apresentação do sistema preditivo (algo em torno de 4min - 10min). **Não se esqueça que tanto a visão do sistema preditivo quanto o dashboard analítico deve ser apresentado em uma visão de negócio.**

Lembre-se de que você poderá apresentar o desenvolvimento do seu projeto durante as lives com docentes. Essa é uma boa oportunidade para discutir sobre

as dificuldades encontradas e 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! 