



Data Scientist Technical Challenge

Contexto

A fraude em plataformas de pagamentos é uma problemática mundial que representa perdas econômicas em grande escala. O MercadoPago, dado a sua dimensão e presença multinacional, é alvo desse tipo de ataque diariamente. Construir modelos robustos e escaláveis de machine learning para prevenção de fraudes que analisem todo fluxo de pagamentos é indispensável para melhorar a experiência dos usuários e reduzir perdas financeiras.

Enunciado

Se deseja construir uma solução baseada em Machine Learning para **predizer se uma transação é fraudulenta ou não**, com base em um conjunto de dados proveniente do sistema de prevenção de fraudes do MercadoLivre.

A fazer:

Resolver o problema de **classificação de transações fraudulentas / não fraudulentas**. Sabemos que para cada transação que é aprovada, temos 25% de lucro. Quando há fraudes aprovadas perde-se 100% do dinheiro da transação. Realizar uma análise e **determinar um modelo que permita maximizar o lucro da empresa**.

1. Desenvolver os seguintes pontos para abordar a solução proposta:
 - Hipótese: Introdução e desenvolvimento.
 - Análise do dataset e transformações dos dados
 - Modelo(s) utilizado(s)
 - Avaliação
 - Conclusão
2. Disponibilizar o código fonte utilizado para a análise no Github ou [Colab](#), junto com os passos necessários para sua execução. Tenha em conta que:
 - A qualidade do código será levada em conta para a avaliação. Recomendamos que escreva um código prolixo e comentado. Esperamos que se expresse o melhor possível para explicar suas decisões e respostas, deixando explícitas as suas suposições.
 - O código deve poder ser executado pelos revisores seguindo os passos indicados, obtendo-se resultados similares ou iguais aos indicados na avaliação e conclusão.

O dataset "MercadoLivre Inc. Data Scientist Hiring Test - Fraud Dataset - Data.csv" se encontra em conjunto com o e-mail presente.

ESCLARECIMENTO: O formato dos números encontra-se com o delimitador . (ponto) para separação decimal e , (vírgula) para milhares.