

## Prevendo Transações dos Clientes

### 1 Introdução

O objetivo deste projeto é identificar quais clientes de uma instituição financeira irão executar uma determinada transação no futuro. Os dados disponíveis para este projeto são dados reais de uma instituição financeira. No entanto, por questões de sigilo, todos os dados sensíveis foram anonimizados.

### 2 Dados

Para a execução deste projeto você deverá utilizar três arquivos CSV:

- arquivo com variáveis independentes para treinar o modelo: data/transacoes\_financeiras\_x\_treinamento.csv
- arquivo com variável dependente para treinar o modelo: data/transacoes\_financeiras\_y\_treinamento.csv
- arquivo de teste com as variáveis independentes: data/transacoes\_financeiras\_x\_teste.csv

### 3 Método de trabalho

Sugere-se seguir as seguintes etapas para a execução desta análise:

- Aquisição, análise exploratória e pré-processamento dos dados;
- Construção do modelo;
- Análise do modelo;
- Preparação do relatório com as estimativas de resultados, e;
- Geração do arquivo com os resultados para os exemplos no conjunto de teste cego (validação final).

### 4 Itens que devem ser entregues

Os seguintes itens devem ser entregues:

- Um relatório apresentando os principais resultados da análise. Este relatório deve ser entregue na forma de um Notebook Python.
- O conteúdo do mesmo deve possuir os seguintes itens:
  - introdução, com descrição do contexto e objetivo do trabalho;
  - descrição do método utilizado para alcançar o objetivo, e;
  - conclusão, ou seja, descrição dos resultados encontrados com a análise realizada.

- O relatório deve apresentar os valores de F1-Score obtidos pelos modelos testados. Além de apresentar os valores de F1-Score, a equipe deverá explicar qual é a equação e o conceito por trás desta equação. Além disso, deve também explicar o motivo de utilizar F1-Score e não acurácia.
- Todos os comandos que fazem parte do script (aquisição, transformação, criação do modelo, avaliação, etc) devem estar presentes no Notebook Python.
- Arquivo texto com as predições para o arquivo de teste (data/transacoes\_financeiras\_y\_teste.csv)

## 5 Equipes e prazos

Este trabalho poderá ser feito em equipes com até três (3) integrantes e deverá ser apresentado no dia 23 de outubro de 2019.

## 6 Avaliação

A avaliação por equipe será composta pelos seguintes atributos:

1. Relatório (60% nota): o relatório precisa estar completo e correto. Além disso, precisa fazer uso de comentários em markdown para facilitar a leitura. Deve ser objetivo e claro.
2. Apresentação (20% da nota): a equipe deve apresentar os resultados de forma clara e objetiva, fazendo uso do python notebook para discutir os resultados com os demais alunos da disciplina.
3. Resultado na forma do arquivo com as predições (20% da nota):

$$nota\_predicoes = F1\_score * 0.2 \quad (1)$$

A nota final do relatório é calculado por:

$$nota\_final = \frac{6 * relatorio + 2 * apresentacao + 2 * F1\_Score}{10} \quad (2)$$