

# Plano de Avaliação - Mentoria de Python

AI for Customer

## Descrição da atividade

### Avaliação

Serão avaliados diferentes aspectos neste projeto: tanto a capacidade de apresentar e interpretar informações, como o pré-processamento, treinamento e avaliação dos modelos.

### Etapa de testes com diferentes Modelos (90 pontos)

Devem ser realizados testes iniciais considerando diferentes modelos de classificação para determinar o mais adequado para o problema em questão.

Quesitos que serão considerados para compor a pontuação total:

- Pré-processamento dos dados (30 pontos)
  - Filtragem do texto (remoção de pontuação, tratamento de flexão, etc.)
  - Tokenização
  - Testes com remoção de *stopwords* e lematização
  - Visualização dos dados (por exemplo, histograma de palavras, *wordcloud*, histograma com a quantidade de textos por classe, etc.)
  - Uso de word embeddings (BoW, TF-IDF, ou no caso do *tensorflow* camadas de Embedding)
  - Normalização dos dados
- Testes e comparação entre modelos (60 pontos)
  - Modelos escolhidos
  - Execução de Fine Tuning nos modelos
  - Cálculo e escolha das métricas

### Apresentação (10 pontos)

Deverá ser feita uma apresentação, mostrando o passo a passo da solução. Desde o tratamento dos dados até a seleção do melhor modelo.

### Dicas:

- 1) Durante a etapa de testes é sugerido usar como modelo de *baseline* a Regressão Logística.
- 2) Para gerar nuvens de palavras é sugerido o pacote *wordcloud*.