

Professor: Samuel Martins (samuel.martins@ifsp.edu.br)

Projeto em Grupo

1. Especificação

Nesta atividade, cada grupo (no máximo 3 integrantes) deverá desenvolver uma solução para automatizar alguma tarefa de *Natural Language Processing*. **O tema é livre!** Quanto mais inovador e/ou desafiador for o problema, melhor avaliado será o projeto. Portanto, **me surpreendam!**

O projeto será dividido em duas partes, sendo que os pontos gerais a serem avaliados em cada parte incluem:

Parte 1:

- Descrição e motivação do problema;
- Descrição da base de dados (e do processo de obtenção dos dados, se for o caso)
- Objetivo de negócio ou científico associado ao problema
- Etapa de limpeza e pré-processamento

Parte 2:

- Extração de características
- Modelos de Machine Learning
- Protocolo de experimentos e validação
- Discussão dos resultados e trabalhos futuros
- **Deployment da solução em produção***
 - A solução precisará estar em produção e que seja de fácil acesso e uso.
 - O grupo deverá investigar por alternativas para isso.
- **Vídeo de até 10 min explicando o projeto!**
 - Todos os membros precisarão participar.

2. Submissão

A submissão desta atividade será em duas partes em tarefas específicas no Moodle da disciplina. O grupo deverá manter o projeto em um repositório no GitHub, cujo link deverá ser disponibilizado para a avaliação. O README do repositório deve ser claro e instruir os arquivos para avaliação (código, relatórios, jupyter-notebook, etc). O vídeo de apresentação do projeto deverá ser disponibilizado na parte 2 para avaliação.

Parte 1 – deadline: 16/04/2023

Parte 2 – deadline: 04/06/2023