

# MBA EM INTELIGÊNCIA ARTIFICIAL E BIG DATA

- PROBLEMAS -> INSERIR FIGURAS NO LATEX -> DISPOSIÇÃO DO TEXTO
- MOTIVAÇÃO BEM ESTRUTURADO -> FONTE
- APRESENTAR OS ALGORITMOS -> NECESSÁRIO
- DIVIDIR O TRABALHO EM VÁRIOS -> MELHOR
- REVISAR A PARTE TEXTUAL -> NOTEBOOKS
- PEDIR AJUDA PARA TERMINAR O TCC

## Parte 1

### PROBLEMAS / CONTEXTO

HIPÓTESES

MERCADO QUE ATUA

INDÚSTRIA DOS GAMES

SERIES

MENTORIA DO CURSO DA USP

AUXILIAR NO DESENVOLVIMENTO DO PROJETO

EMPRESAS DE STREAMING -> MAX, APPLE TV +

REDES SOCIAIS -> TWITTER, NETFLIX, DISNEY +, YOUTUBE

## Parte 3

### PRÉ-PROCESSAMENTO

- LIMPEZA DOS DADOS -

LIDAR COM PALAVRAS EM INGLÊS -> PORTUGUÊS

DESCOBRIR QUANTIDADE -> COLETAR COMENTÁRIOS

VERIFICAR A TRADUÇÃO -> VISUALMENTE

- EMOJI -> NOTAS DE RODAPÉ NO TRABALHO

-> QUANTIFICAR NO TEXTO

-> TRANSFORMAR EM TEXTO

- ERROS DE PORTUGUÊS -> LIDAR

= PALAVRAS EM INGLÊS

-> FRASE COMPLETA E/OU PARTE DO TEXTO

TECNICAS DE ETL

TABELA DE TRADUÇÃO -> EMOJI -> SIGNIFICADO

DISPONIBILIZAR

-> NO TRABALHO

## Parte 2

### COLETA

ESCOLHER A FONTE DE DADOS

METODO DE COLETA -> API

-> WEBSERPING

LINGUAGEM DE PROGRAMAÇÃO -> PYTHON

DICIONÁRIO DOS DADOS

DESCRIÇÃO

ARMAZENAMENTO -> INTERNO

-> EXTERNO

BANCO DE DADOS -> CLOUD

PLANILHA, NOSQL, COMPUTING

-> CSV

VERSIONAMENTO -> GITHUB

ANOTAÇÕES -> GOOGLE COLAB, GOOGLE DRIVE

-> ARTIGO -> MEDIUM

-> KAGGLE

COLETA ORIENTADA AOS OBJETIVOS

ESCALAR

ENRIQUECIMENTO DOS DADOS

EXISTEM BIBLIOTECAS PARA CORRIGIR

-> DIFERENCIAR LINGUAGEM FORMAL

-> LINGUAGEM DA INTERNET

-> GÍTIAS

USO DA LMI -> PARA DETECTAR

-> NO TEXTO