Guia 1

| **COMPONENTE CURRICULAR:** | **Projeto Aplicado II** |
| --- | --- |
| **GRUPO** | **Grupo - Projeto Aplicado 13** |
| **ALUNOS** | |
| **Thaís Cristine de Andrade Gomes** | **10721642** |
| **Paulo Ricardo de Oliveira Ramos** | **10721464** |
| **Lucas Iglezias dos Anjos** | **10433522** |
| **Oscar Augusto de Oliveira Luz** | **10435099** |

**Análise de Avaliações de Usuários do Aplicativo Tinder**

| **Sumário** |
| --- |

1. Premissas do Projeto ................................................................................................. 3
2. Objetivos e Metas ....................................................................................................... 4
3. Cronograma de Atividades (Estimativa) ..................................................................... 4
4. Link do Projeto no GitHub .......................................................................................... 5

| **Premissas do Projeto** |
| --- |

**1.1 Definição da Organização**

A organização escolhida para o desenvolvimento do projeto é o Tinder, um aplicativo de relacionamento conhecido e utilizado mundialmente. Fundado em 2012, o Tinder conecta pessoas com base em preferências pessoais e localização geográfica, sendo reconhecido por sua interface intuitiva de interação por "match".

A empresa atua na área de tecnologia, com foco em soluções digitais voltadas para a conexão entre pessoas. Sua principal plataforma é o aplicativo móvel, disponível na Google Play Store e na App Store.

**1.2 Apresentação dos Dados (Metadados)**

A base de dados utilizada neste projeto consiste em avaliações públicas feitas por usuários do aplicativo Tinder na Google Play Store. O arquivo, no formato CSV, foi obtido por meio da plataforma Kaggle, por meio do seguinte repositório: <https://www.kaggle.com/datasets/shivkumarganesh/tinder-google-play-store-review>.

O conjunto de dados contém 676.014 registros com avaliações textuais, classificações de 1 a 5 estrelas, data da avaliação, nome do usuário, número de curtidas na avaliação e eventuais respostas fornecidas pela empresa. Os principais campos utilizados na análise são:

* *review*: texto da avaliação fornecida pelo usuário
* *rating*: nota atribuída ao aplicativo (variando de 1 a 5)
* *review\_date*: data da publicação
* *review\_reply*: resposta oficial da empresa (se houver)
* *review\_likes*: número de curtidas da avaliação

A análise será conduzida com base nas avaliações textuais dos usuários, aplicando-se técnicas de ciência de dados para classificar os sentimentos expressos por eles. O projeto tem como propósito simular uma aplicação prática de monitoramento automatizado da experiência do usuário.

| **Objetivos e Metas** |
| --- |

**2.1 Objetivo Geral**

Analisar as avaliações dos usuários do aplicativo Tinder publicadas na Google Play Store, aplicando técnicas de ciência de dados e processamento de linguagem natural (PLN) para identificar padrões de sentimento, temas recorrentes e oportunidades de melhoria.

**2.2 Metas Específicas**

1. Explorar a base de dados com foco nos campos textuais e numéricos.
2. Realizar o pré-processamento das avaliações textuais, incluindo limpeza e tokenização.
3. Agrupar avaliações conforme a nota atribuída, estabelecendo categorias de sentimento.
4. Aplicar modelos de classificação supervisionada para prever o sentimento com base no texto.
5. Gerar visualizações que permitam identificar temas recorrentes nas avaliações.
6. Consolidar os resultados em formato de storytelling, relacionando-os a ações que poderiam ser adotadas pela equipe de produto do Tinder.

| **Cronograma de Atividades (Estimativa)** |
| --- |

O cronograma abaixo representa uma estimativa de esforço por semana e poderá sofrer ajustes conforme a evolução do trabalho em grupo.

| **Semana** | **Atividade** | **Prazo** |
| --- | --- | --- |
| 1 | Definição do grupo, empresa e área de atuação | 2 dias |
| 1 | Análise inicial do dataset | 2 dias |
| 1 | Redação das premissas do projeto e repositório inicial | 2 dias |
| 2 | Análise exploratória dos dados | 5 dias |
| 2 | Pré-processamento dos textos | 4 dias |
| 3 | Treinamento do modelo de classificação | 6 dias |
| 3 | Visualizações e gráficos | 3 dias |
| 4 | Redação do storytelling com os resultados | 3 dias |
| 4 | Organização da entrega final no GitHub | 3 dias |

| **Link do Projeto no GitHub** |
| --- |

O repositório do projeto foi desenvolvido para centralizar todos os materiais produzidos ao longo do semestre do Projeto Aplicado II. O link é:.