

Projeto Final: Análise de Avaliações de Consumidores

Introdução à Ciência de Dados

Entrega: 25/05

Objetivo

Analisar avaliações de consumidores para diversos produtos ou serviços para identificar padrões de satisfação do cliente e áreas para melhoria, utilizando técnicas de raspagem de dados web, análises no Excel e insights potencializados por IA.

Visão Geral do Projeto

O grupo deve extrair dados de avaliações de consumidores de uma categoria de produto ou serviço, utilizar a API da OpenAI para obter insights textuais mais profundos e analisar os dados no Excel.

Todas as funcionalidades devem ser divididas em módulos, e a organização das funções será analisada. Apenas código documentado será avaliado, e todo o código deve seguir o mesmo padrão e observar boas práticas de nomeação de identificadores discutidas em sala.

A entrega será realizada por e-mail, com todos os participantes em cópia. O texto do e-mail deve conter o nome completo dos alunos e o link para um repositório público do GitHub. É esperado que o trabalho seja desenvolvido em pequenos incrementos, e a evidência deste fluxo de trabalho é o histórico de commits do repositório.

Etapas Detalhadas

1. Extração dos Dados

- O grupo deve selecionar um site compatível com extração de dados e uma categoria de produto ou serviço para analisar.
- O grupo deve coletar dados de avaliação, classificação, depoimentos e informações do usuário, para citar alguns, e armazená-los em estruturas de dados nativas de Python, como listas e dicionários.

2. Preparação e Limpeza de Dados

- O grupo deve limpar e organizar os dados utilizando Python, focando na remoção de duplicatas, tratamento de valores ausentes e diferentes formatações.
- [EXTRA] O grupo pode, se desejar, criar características adicionais na base de dados, documentando a justificativa.

3. Geração de Insights Usando a API da OpenAI

- O grupo deve utilizar a API da OpenAI para processar os dados textuais extraídos. O objetivo desta etapa é se apoiar em LLMs para potencializar a capacidade de análise do trabalho, como resumir e interpretar avaliações, agrupar temas ou traduzir conteúdo.

4. Análise de Dados Usando o Excel

- O grupo deve exportar os dados coletados e enriquecidos para um arquivo Excel.
- O grupo deve elaborar 3 peças de análise – sejam tabelas dinâmicas, gráficos ou conjuntos de fórmulas – para avaliar aspectos como a distribuição das classificações, padrões de frequência, reclamação ou elogios.

5. Objetivos de Aprendizagem

Este projeto tem como objetivo inserir os alunos em um problema de dados comum no mundo corporativo, fomentando a aplicação das habilidades desenvolvidas no primeiro período. É esperado que o grupo demonstre:

- Domínio da extração de dados da web para obtenção de informação sobre consumidores.
- Familiaridade de uso de IA para aprimorar a análises de dados.
- Utilização de técnicas avançadas de Excel para análise de dados de negócio.