

# Projeto de Formatura – 2023



## PCS - Departamento de Engenharia de Computação e Sistemas Digitais

Engenharia de Computação

Tema: **Análise de Sentimento e Dashboard de preços de Produtos Obtidos por Web Scraping**

### Introdução

Vivemos em um contexto no qual Big Data e Machine Learning estão cada vez mais em alta. Cada vez mais, temos dados sobre produtos, demanda e feedback dos usuários disponíveis na Web. Cada e-commerce possui milhões de dados, e temos milhares de e-commerce espalhados pela web, muitas vezes com produtos semelhantes. Porém tais dados públicos estão espalhados, não estruturados e não são facilmente acessados. Eles não servem de nada nem criam valor se não forem devidamente utilizados. Nosso projeto se propõe em transformar esses dados em informação e inteligência.

### Objetivo

Dar meios para que uma empresa avalie estrategicamente seus produtos e de seus concorrentes, analisando os preços e sentimentos de seus consumidores.

### Produto

Disponibilizar para o nosso cliente a informação de qual é o preço de um conjunto de produtos alvo, como tais preços variaram e qual o sentimento dos compradores de acordo com cada um desses produtos. Fazendo esse acompanhamento constante podemos ver tanto como a demanda, a oferta e o feedback dos compradores se comportou. Disponibilizamos tais informações em DashBoard, com sugestões de tomada de decisão; mas também possibilitando que os clientes criem sua estratégia de negócio baseada nos dados.

### Resultado

Podemos verificar que nosso projeto cumpre o seu objetivo, e agrega valor. Testamos com um nicho voltados pra farmácias (dado a quantidade de farmácias familiares versus grandes redes), e o acompanhamento dos dados e dos comentários se mostrou relevante e promissor. Os dashboards foram claros e permitiram insights positivos.

**Integrantes:**

- Andrei dos Santos
- Henrique Geribello Giabardo
- Lucas Lopes de Paula Junior

**Professor Orientador:** Prof. Dr. Jorge Becerra Risco

### Motivação

A importância de acompanhar a concorrência, os preços gerais dos produtos e sua demanda é clara para qualquer comerciante. Porém, são poucos hoje em dia que tem a tecnologia basal para fazer esse acompanhamento. Grandes players e empresas possuem os recursos e know-how para contratar os melhores desenvolvedores para acompanhar tais dados, porém competidores menores não têm tal estrutura técnica. Nosso projeto se propõe a equilibrar o jogo, disponibilizando essas informações, e permitindo que mais negócios possam tomar decisões inteligentes baseadas nos dados.

### Arquitetura

O projeto se destaca por três pilares fundamentais:

- Web Scraping Avançado ( em Python): O grupo utilizou o framework Scrapy para criar um sistema de coleta de dados (preços) no nicho de farmacia. Esse sistema faz ETL (Extract Transform Load) dos dados, fazendo a extração de tais dados da Web de sites previamente escolhidos, transformando os dados para deixá-los de maneira estruturada e armazenando numa base própria.
- IA para Análise de sentimento: Utilizamos a API do google de Análise de Sentimentos (Cloud Natural Language) para processar os comentários armazenados pelo subsistema anterior, avaliando o grau de satisfação do consumidor a respeito de aspectos do produto. Assim fornecemos uma compreensão mais quantificada do feedback dos compradores.
- Dashboard de Dados e Insights: Utilizamos ferramentas Open Source para demonstrar os dados, bem como os Insights e sugestões gerados.

WebScrapper/Sentiment Analysis Component Diagram

