

## **Estudo de Caso 3**

**Contexto:** Seja bem-vindo ao projeto sobre Sistemas de Recomendação. Trabalharemos com o conjunto de dados de análises de produtos da Amazon para este projeto. O conjunto de dados contém classificações de diferentes produtos eletrônicos. **Amazon electronics Rating Dataset** (<https://www.kaggle.com/datasets/vibivij/amazon-electronics-rating-datasetrecommndation/data>).

**Objetivo:** Você é Cientista de Dados na Amazon e recebeu a tarefa de construir um sistema de recomendação para recomendar produtos aos clientes com base em suas classificações anteriores para outros produtos. Você tem uma coleção de dados rotulados de análises de produtos da Amazon. O objetivo é extrair insights significativos dos dados e construir um sistema de recomendação que ajude a recomendar produtos aos consumidores online.

O conjunto de dados da Amazon contém os seguintes atributos:

**userId:** cada usuário identificado com um ID exclusivo

**productId:** cada produto identificado com um ID exclusivo

**Avaliação:** A avaliação do produto correspondente pelo usuário correspondente

**Timestamp:** Hora da classificação. Não usaremos esta coluna para resolver o problema atual

1) **Na primeira parte desta atividade** você deve criar o mecanismo de recomendação de produtos da Amazon citado acima, utilizando o algoritmo **KNN**. Você pode se inspirar no notebook que se encontra no Moodle na pasta:

- **Aula11 ->**

**SistemaRecomendacaoPythonCEUB KNN Final.ipynb.**

Ele é um mecanismo de recomendação que utiliza o algoritmo **KNN**, para sugerir filmes aos usuários, no estilo Netflix.

2) **Na segunda parte desta atividade** você vai desenvolver o mesmo mecanismo de recomendação de produtos da Amazon, mas desta vez utilizando o algoritmo **k-means**, com clusterização. Você pode se inspirar no notebook disponível no Kaggle:

- **Online Retail Clustering Recommendation System**

(<https://www.kaggle.com/code/samantas2020/online-retail-clustering-recommendation-system>).

Você também pode se inspirar no notebook que se encontra no Moodle na pasta:

- **Aula 12 ->**

**SistemaRecomendacaoPythonCEUB\_KMeans.**

- a) Lembre-se de tanto na primeira parte quanto na segunda fornecer explicações detalhadas a respeito de cada etapa;
- b) Explique a diferença entre as duas abordagens (KNN e K-Means);
- c) Explique qual abordagem apresentou os melhores resultados e porquê.