

Lista de Exercícios 10

(prazo final para entrega: 27/06/2019 - segunda)

Suponha que a internet caiu no Campus Quixadá e você quer fazer uma consulta através de palavras chave sobre as reviews de filmes do IMDB.

- 1) Defina sua consulta (palavras chave)**
 - 2) Consulte o dataset**
 - 3) Mostre as 10 reviews mais semelhantes à sua consulta.**
- Dica: Vetorize as reviews usando CountVectorizer ou TF-IDF e raqueie a similaridade usando a distância cosseno ou Jaccard.
 - Informações sobre o dataset e como transformá-lo para uso, podem ser obtidas no [seguinte link](#).
 - Há um csv compactado já com as reviews prontas para uso no [link](#). O read_csv do pandas consegue lê-lo diretamente. Você nem precisa descompactá-lo antes.
 - Veja também os jupyter notebooks que fizemos esta semana nas aulas de Laboratório.