





# Exercícios – DM

## Filtragem Colaborativa

Prof. Renê R. Veloso

- 1) Considere a seguinte tabela:

Artistas:	Ana	Axel	Bruno	Daniel	Jessica	Larissa	Nilton	Yuri
Muddy W.	3.5	2	5	3	-	-	5	3
Junior Wells	2	3.5	1	4	4	4.5	2	-
Bonnie Lee	-	4	1	4.5	1	4	-	-
Eric Clapton	4.5	-	3	-	4	5	3	5
Gary Smith	5	2	5	3	-	5	5	3
B.B. King	1.5	3.5	1	4.5	-	4.5	4	2.5
James Cotton	2.5	-	-	4	4	4	5	3
Robert J.	2	3	-	2	1	4	-	-

- Calcule a distância (Manhattan) entre Jéssica e Ana.
  - Quais é o problema em usarmos medidas de distância simples como Manhattan e Euclidiana nesse conjunto de dados?
  - Qual é a métrica de distância ideal para esse conjunto de dados? Por quê?
  - Calcule a pessoa mais próxima do Yuri usando a métrica ideal. (mostre os cálculos)
  - Usando o KNN, com  $k=3$ , a Jéssica iria gostar de escutar B. B. King? (mostre as etapas do processo com os cálculos)
  - Qual seria a recomendação para uma pessoa que escutou o Gary Smith?
- 2) Quais são os problemas geralmente encontrados quando se pensa em construir um sistema de recomendação?
- 3) Suponha que você está construindo um sistema onde os usuários avaliam os itens usando  ou . Cada item  $i$  possui um vetor de características  $v_i = \langle a, b, c, d, e, f \rangle$ , onde  $a, b, c, d, e, f$  são valores entre 0 e 1. Dado um novo item  $j$  com vetor de características  $v_j = \langle x, y, z, w, k, p \rangle$ , escreva, em pseudo-linguagem, como encontrar qual seria a classificação ( ou ) desse novo item para um determinado usuário  $u$ .
- 4) Há um dataset muito conhecido na Web, o MovieLens, que possui avaliações de vários filmes, encontrado no endereço <http://grouplens.org/datasets/movielens/>.

- a. Escreva um recomendador UbF e verifique quais são as recomendações para o usuário 87.
  - b. Sabendo-se que o usuário 87 deu nota 5 para o filme “Star Trek: The Wrath of Khan (1982)”, quais outros filmes podem ser recomendados a ele?
- 5) Faça o download do dataset sobre avaliações de cervejas em [https://s3.amazonaws.com/demo-datasets/beer\\_reviews.tar.gz](https://s3.amazonaws.com/demo-datasets/beer_reviews.tar.gz)
  - a. Quais são as top-5 cervejas mais similares a Coors Light?
  - b. Quais são as top-5 cervejas mais similares a Stella Artois?
  - c. Quais são as top-5 cervejas mais similares a Brahma?
- 6) Utilize a biblioteca Python para sistemas de recomendação disponível em <http://muricoca.github.io/crab/tutorial.html> e compare com os resultados dos exercícios 4 e 5.
  - a. Alternativa para programadores Java: <http://lenskit.grouplens.org/>