

LISTA SOBRE PREVISÃO E RECOMENDAÇÃO

Prof. Filipe Braida do Carmo

Introdução à Teoria de Sistemas de Recomendação (TM429)

Os exercícios deverão utilizar o *dataset* MovieLens [1]. Essa base de dados vem de um sistema de recomendação de filmes desenvolvido pelo GroupLens e possui 100.000 avaliações. Nela existem 943 usuários e 1682 filmes, e o seu conjunto de preferência é um número inteiro que varia entre um a cinco. Além das avaliações, existem informações sobre usuários, como a idade e o sexo, e sobre os filmes, tais como título do filme e a data de lançamento do filme. Essas informações não serão utilizadas nos exercícios dessa lista.

Questão 1: Gere a lista dos dez filmes mais bem avaliados do conjunto de dados utilizando a média simples.

Questão 2: Gere a mesma lista do exercício anterior, mas utilizando a *Damped Mean* com k igual a 30. Faça uma comparação com o exercício anterior.

Questão 3: Considere que o usuário goste de um filme quando a nota que ele atribui a um filme seja superior ou igual a 4 e não goste caso contrário. Faça uma lista de recomendações dos filmes com o maior percentual positivo de notas. Faça uma comparação com os exercícios anteriores.

Questão 4: Considerando a mesma condição do exercício anterior, faça uma lista de recomendações considerando os filmes que possuem mais notas iguais ou superiores a 4. Faça uma comparação com os exercícios anteriores.

Questão 5: Considerando a mesma condição do exercício anterior, faça uma lista de recomendações considerando os filmes que possuem mais notas iguais ou superiores a 4 do que notas 1, 2 e 3. Faça uma comparação com os exercícios anteriores.

Questão 6: Escolha três filmes com distintas avaliações e calcule o histograma de notas recebidas pelos usuários. Faça uma comparação entre os filmes escolhidos.

References

- [1] BN Miller, Istvan Albert, and SK Lam. MovieLens unplugged: experiences with an occasionally connected recommender system. *Proceedings of the 8th ...*, pages 263–266, 2003. URL <http://dl.acm.org/citation.cfm?id=604094>.