



Universidade Federal de Minas Gerais

Recommender Systems - RC2

Tomas Lacerda Muniz e Lorenzo Carneiro Magalhães
2021088116 2021031505

Sumário

1	Introdução	2
2	Metodologia	2
3	Conclusão	2

1 Introdução

Este projeto prático visa desenvolver um sistema de recomendação de filmes baseado em conteúdo, utilizando metadados de filmes obtidos do OMDb. Conforme discutido em sala de aula, várias escolhas de implementação afetam a qualidade das recomendações baseadas em conteúdo. Isso inclui opções para representação de conteúdo (por exemplo, unigramas, n-gramas, conceitos, entidades nomeadas, tópicos latentes), perfilamento do usuário (por exemplo, agregando feedbacks positivos e negativos), e estimativa de similaridade entre usuário e item. Neste documento, serão exploradas a nossa implementação desses componentes, com o objetivo de verificar o desempenho das recomendações produzidas.

2 Metodologia

1. Configuração Inicial e Importação de Dados:

- Importa bibliotecas necessárias como `os`, `pandas`, `numpy`, `sys` e `surprise`.
- Define variáveis globais como `test_size` e `random_state`.
- Lê os argumentos da linha de comando para obter os caminhos dos arquivos de classificações (`ratings_file`), conteúdo (`content_file`) e alvos (`targets_file`).

2. Carregamento e Preparação dos Dados:

- Carrega os dados de classificações e conteúdo usando funções do módulo `data_loader`.
- Divide os dados em conjuntos de treino e teste, com a possibilidade de usar todo o conjunto de dados para treinamento.

3. Construção e Treinamento do Modelo:

- Cria uma instância do modelo SVD a partir do módulo `SVD_model`.
- Treina o modelo com os dados de treino. Se um modelo pré-treinado estiver disponível, ele pode ser carregado em vez de treinar um novo.

4. Avaliação do Modelo:

- Se o conjunto de teste for definido, o modelo é avaliado usando métricas como RMSE (Root Mean Square Error) e MAE (Mean Absolute Error).

5. Geração de Previsões para Alvos:

- Carrega um conjunto de dados de alvos e gera previsões para eles.
- Utiliza informações adicionais dos itens para calcular um rating final ponderado.

6. Processamento e Salvamento dos Resultados:

- Normaliza as classificações previstas e as salva em arquivos CSV, juntamente com os IDs dos usuários e itens recomendados.

3 Conclusão

Nós observamos ao longo do trabalho que as recomendações não personalizadas baseadas em popularidade dos itens eram extremamente poderosas, usando então nosso SVD para gerar uma certa personalização conseguimos ultrapassar todos os parâmetros necessários pré estabelecidos.

Uma reflexão que sobra deste trabalho é que as pessoas tem sim sua particularidade mas que obras bem feitas muitas vezes sobresaem os gostos individuais, se tornando gosto da maioria e algo fácil e poderoso de se recomendar.