

Sistema de Recomendação de Filmes Baseado no Algoritmo Apriori

Giovanna Salomão Rodrigues

Samira de Jesus Santos

Victor Rogério Aguiar do Rosário

Universidade / Curso: Sistemas de Informação

Email: giovanna@email.com

Resumo

Este artigo descreve o desenvolvimento de um sistema de recomendação de filmes para a plataforma fictícia Manga Play, com base no algoritmo Apriori. O sistema recomenda novos títulos com base nas avaliações anteriores dos usuários, considerando como 'gostou' as notas maiores que 3. A solução foi desenvolvida sem o uso de bibliotecas prontas para o algoritmo, respeitando as restrições do desafio. O projeto contempla a construção do modelo de recomendação baseado em regras de associação e sua aplicação tanto sobre o histórico de avaliações quanto sobre o último filme bem avaliado pelo usuário.

1. Introdução

O consumo de conteúdo audiovisual tem se transformado com os serviços de streaming. Entretanto, a dificuldade de encontrar novos títulos semelhantes aos já assistidos motivou o desenvolvimento de sistemas de recomendação. Este trabalho propõe um sistema baseado no algoritmo Apriori para sugerir filmes com base nas preferências dos usuários da plataforma fictícia Manga Play.

2. Fundamentação Teórica

O algoritmo Apriori é um dos mais utilizados para a descoberta de regras de associação. Ele parte do princípio de que um subconjunto de um itemset frequente também deve ser frequente. As principais métricas utilizadas são o suporte (frequência relativa de ocorrência), a confiança (probabilidade condicional de ocorrer o consequente dado o antecedente) e o lift (relação entre confiança e frequência esperada).

Pseudocódigo do Apriori:

1. Gerar todos os conjuntos de itens frequentes com suporte mínimo;
2. Para cada conjunto frequente, gerar regras possíveis com confiança mínima;
3. Calcular métricas e ordenar regras para recomendação.

3. Metodologia

Inicialmente, foram utilizados dois datasets: um catálogo de filmes (movies.csv) e um histórico de avaliações de usuários (ratings.csv). Ambos foram unificados usando Pandas para relacionar filmes e avaliações.

No pré-processamento, foi definido que usuários gostaram de um filme se atribuíram nota superior a 3. A partir disso, foram criadas transações contendo os filmes curtidos por usuário.

O algoritmo Apriori foi implementado em Python sem o auxílio de bibliotecas específicas de mineração, seguindo os passos descritos anteriormente. As métricas mínimas adotadas foram: suporte mínimo (0.020.020.02) e confiança mínima (0.30.30.3).

O sistema implementado recomenda filmes baseando-se no histórico completo de filmes curtidos pelo usuário.

4. Resultados e Discussões

A aplicação do algoritmo resultou em diversas regras de associação relevantes. Por exemplo, uma das regras obtidas foi:

$$\{A \text{ Origem}\} > \{\text{Interestelar}\}$$

com confiança aproximada de 0.450.450.45 e suporte 0.040.040.04.

A definição dos parâmetros (suporte 0.020.020.02, confiança 0.30.30.3) mostrou-se adequada para identificar relações relevantes sem gerar excesso de regras triviais.

A abordagem adotada demonstrou que o sistema é capaz de fornecer recomendações coerentes baseadas em padrões frequentes de preferência.

5. Considerações Finais

O trabalho resultou em um sistema funcional que fornece recomendações relevantes baseadas no histórico de preferências. Pontos positivos incluem a flexibilidade na definição de métricas e independência de bibliotecas externas para o Apriori.

Como limitações, o algoritmo Apriori pode enfrentar dificuldades de desempenho em bases maiores, dada a explosão combinatória de candidatos.

Em trabalhos futuros, sugere-se a incorporação de métricas adicionais (como lift) para refinar recomendações e integração com técnicas híbridas que combinem filtragem colaborativa.

Referências

- [1] Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules. In Proc. 20th Int. Conf. Very Large Data Bases (VLDB), 487–499.
- [2] Tan, P.-N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining. Pearson.
- [3] SBC. Modelo para Publicação de Artigos. <https://www.sbc.org.br/>