

HistFlix: Um Sistema Personalizado para Filmes Históricos e Documentários

Bruno Baltuilhe, Isaque Pimentel, Kelly Graziely Pena

Universidade Presbiteriana Mackenzie

May 14, 2025

HistFlix: Um sistema de recomendação projetado para sugerir filmes e documentários históricos com base nas preferências e emoções dos usuários.

- ▶ Combina filtragem colaborativa (FC) e filtragem baseada em conteúdo (FBC).
- ▶ Integra análise de sentimentos para personalizar recomendações.
- ▶ Objetiva melhorar o aprendizado e o engajamento com conteúdos históricos.

Desafios em Sistemas de Recomendação:

- ▶ Foco educacional: Necessidade de um sistema voltado para conteúdos históricos e educacionais (ODS da ONU).
- ▶ Personalização limitada: Recomendações genéricas não capturam emoções dos usuários.
- ▶ Problema do novo usuário: Falta de dados para novos usuários ou itens.

Objetivo: Desenvolver um sistema de recomendação híbrido que combine filtragem colaborativa, baseada em conteúdo e análise de sentimentos.

Conjunto de Dados MovieLens 1M:

- ▶ 1 milhão de avaliações de 6.000 usuários em 4.000 filmes.
- ▶ Inclui metadados como gêneros, títulos e anos de lançamento.

Dados Adicionais:

- ▶ Integrado com a API TMDb para obter:
 - ▶ Sinopse dos filmes.
 - ▶ Lista de atores.
 - ▶ Avaliações de usuários.

Sistema de Recomendação Híbrido:

- ▶ **FC:** Utiliza SVD (Decomposição em Valores Singulares) para prever avaliações de usuários com base em usuários semelhantes.
- ▶ **FBC:** Calcula similaridade usando:
 - ▶ Gêneros, títulos e anos de lançamento.
 - ▶ Pontuações de sentimentos das avaliações dos usuários.
- ▶ **Modelo Híbrido:** Combina pontuações colaborativas e baseadas em conteúdo:

$$\text{Pontuação Híbrida} = \alpha \cdot \text{Pontuação FC} + (1 - \alpha) \cdot \text{Pontuação FBC}$$

Próxima Etapa: Análise de Sentimentos

Integração de Sentimentos:

- ▶ Avaliações de usuários analisadas usando TextBlob.
- ▶ Pontuações de polaridade variam de -1 (negativo) a +1 (positivo).
- ▶ Pontuações de sentimentos incluídas no cálculo de similaridade de conteúdo.

Exemplo:

- ▶ Sentimento positivo: Recomendar gêneros inspiradores (e.g., Comédia, Aventura).
- ▶ Sentimento negativo: Recomendar gêneros reflexivos (e.g., Drama, Romance).

Componentes Principais:

- ▶ Pré-processamento e enriquecimento de dados usando API TMDB.
- ▶ Filtragem colaborativa (SVD).
- ▶ Filtragem baseada em conteúdo (matriz de similaridade).
- ▶ Análise de sentimentos (TextBlob).

Métricas de Avaliação:

- ▶ **RMSE:** 0.87 (quanto menor, melhor).
- ▶ **Precisão@10:** 0.82.
- ▶ **Revocação@10:** 0.68.
- ▶ **F1-Score:** 0.74.

Exemplos de Recomendações

Entrada do Usuário: "Estou feliz e quero assistir algo inspirador."

	Título	Gêneros
Principais Recomendações:	Toy Story	Animação
	Jumanji	Aventura
	Sabrina	Comédia

Desafios:

- ▶ Problema do novo usuário para novos itens e usuários.
- ▶ Disponibilidade limitada de avaliações de alta qualidade.

Trabalhos Futuros:

- ▶ Expandir o conjunto de dados para incluir mais filmes e usuários (e.g., MovieLens 32M).
- ▶ Melhorar a análise de sentimentos com modelos avançados de PLN (e.g., BERT).
- ▶ Desenvolver uma interface amigável para recomendações em tempo real.

Principais Conclusões:

- ▶ HistFlix combina filtragem colaborativa, baseada em conteúdo e análise de sentimentos.
- ▶ Oferece recomendações personalizadas e emocionalmente conscientes.
- ▶ Objetiva melhorar o aprendizado e o engajamento com conteúdos históricos.

Obrigado!