



CENTRO UNIVERSITÁRIO DO DISTRITO FEDERAL – UDF

MATHEUS COSTA TORRES

Nº da Matrícula: 24130745

Data: 10/06/2024

Recomenda filmes

Brasília – DF

Documentação do Projeto: Recomenda filmes

1. Descrição do Problema:

O projeto consiste na implementação de um sistema de recomendação de filmes utilizando o algoritmo de inteligência artificial K Nearest Neighbors (KNN). O objetivo é fornecer aos usuários sugestões de filmes com base em suas avaliações anteriores. Isso visa melhorar a experiência do usuário em plataformas de streaming de vídeo, ajudando-os a descobrir novos filmes que possam gostar. O sistema utiliza dados de filmes e avaliações obtidas do Google Sheets.

2. Metodologia Utilizada:

2.1 Coleta de Dados:

Os dados foram coletados a partir de duas planilhas do Google contendo informações sobre filmes e avaliações dos usuários. Foram selecionadas colunas relevantes, como IDs de filmes, títulos, idiomas originais, número de avaliações e IDs de usuários, IDs de filmes e avaliações, respectivamente.

2.2 Pré-processamento dos Dados:

- Os dados foram limpos para remover valores nulos.
- Usuários com menos de 999 avaliações foram filtrados, considerando apenas usuários ativos.
- Filmes com menos de 999 avaliações ou que não estavam em inglês foram excluídos.
- As avaliações duplicadas foram removidas.
- Os IDs de filmes foram convertidos para inteiros.
- Uma tabela pivô foi criada com os títulos dos filmes como índices e os IDs de usuários como colunas.
- A tabela pivô foi convertida em uma matriz esparsa para eficiência computacional.

2.3 Treinamento do Modelo:

- O algoritmo KNN é utilizado para treinar o modelo de recomendação. O modelo é treinado com base nas avaliações dos usuários sobre os filmes. O KNN é uma técnica de aprendizado supervisionado que classifica os pontos de dados com base em sua proximidade com outros pontos de dados no espaço de recursos. No contexto deste projeto, o modelo é treinado para encontrar filmes semelhantes com base nas avaliações dos usuários.

3. Resultados Obtidos:

- Os 10 filmes mais bem avaliados foram listados com suas respectivas médias de avaliação arredondadas para duas casas decimais.
- Após o treinamento do modelo, quando um usuário insere o nome de um filme que gostou, o algoritmo KNN é usado para encontrar os filmes mais próximos em termos de avaliações de usuários. Isso permite que o sistema recomende filmes similares que o usuário provavelmente gostará. Essa é uma aplicação típica de sistemas de recomendação baseados em inteligência artificial, onde o modelo aprende padrões nos dados para fazer previsões ou sugestões personalizadas.

4. Conclusões:

O projeto foi bem-sucedido na implementação de um sistema de recomendação de filmes utilizando inteligência artificial. Os resultados obtidos demonstram a capacidade do modelo de sugerir filmes com base nas avaliações dos usuários. No entanto, há oportunidades para melhorias, como o desenvolvimento de uma interface mais amigável e a incorporação de métricas de avaliação para avaliar a precisão do sistema de recomendação. No geral, o projeto oferece uma base sólida para futuras iterações e melhorias no sistema de recomendação.

Referências

- K-Nearest Neighbors algorithm. Disponível em:
https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm
<https://medium.com/brasil-ai/knn-k-nearest-neighbors-1-e140c82e9c4e>
- Sparse matrix. Disponível em:
<https://www.geeksforgeeks.org/sparse-matrix-representation/>
- Google Sheets API. Disponível em:
<https://developers.google.com/sheets/api/guides/concepts>
- "Como Criar um Sistema de Recomendação de Produtos Usando Machine Learning" por Mário Filho. Disponível em:
<https://mariofilho.com/como-criar-um-sistema-de-recomendacao-de-produtos-usando-machine-learning/>

Observações

- Este trabalho pode ser aprimorado incorporando técnicas de normalização e ponderação de avaliações, além de considerar diversos gêneros cinematográficos para recomendações mais precisas.
- A implementação pode ser expandida para incluir um sistema de feedback do usuário, permitindo que ele avalie as recomendações e refine o modelo ao longo do tempo.