

Projeto Aplicado 3

Membros:

Rafael Castro dos Santos,

Matheus Aparecido Ramos Lourenco,

Gabriel Sicari Moretti.

Professor: Murilo Gleyson Gazzola

Sistema de Recomendação.

25/08/2023

Resumo

Esse documento é a documentação do Projeto Aplicado 3, onde engloba os principais pontos estudados e realizados, desde a escolha da Base de Dados e sua análise, também passando pela organização do projeto como um todo. Não deixando de abordar assuntos um pouco mais técnicos como pacotes e bibliotecas e a escolha do modelo de Machine Learning aplicado. E finalizando com as principais conclusões e pontos observados, e os links que se encontram o projeto de fato.

Sumário

1. Introdução.....	4
1.1 Metodologia.....	4
2. Importando as bibliotecas.....	4
3. Importando os Datasets.....	4
4. Análise exploratória e preparação dos dados.....	5
5. Criação do sistema de recomendação.....	5
6. Exemplos de recomendação.....	5
7. Conclusão.....	5
8. Desafios.....	6
9. Conclusões sobre a pesquisa.....	6
10. Trabalhos futuros.....	6
11. Links úteis.....	6
8.1 Link do Vídeo Youtube.....	6
8.2 Link do GitHub.....	6
12. Referências.....	7

1. Introdução

Os sistemas de recomendação são uma ferramenta importante para auxiliar os usuários na busca por produtos ou serviços que possam ser do seu interesse. No caso dos livros, os sistemas de recomendação podem ajudar os usuários a encontrar novos títulos que sejam compatíveis com seus gostos e preferências.

O objetivo deste projeto é desenvolver um sistema de recomendação de livros baseado em dados. Para isso, será utilizado o dataset "Book-Crossing" disponibilizado no Kaggle. Este dataset contém informações sobre usuários, livros e classificações de livros.

1.1 Metodologia

A metodologia do projeto será dividida em três etapas:

- **Exploração dos dados:** Nesta etapa, serão realizados estudos exploratórios sobre os dados do dataset. O objetivo é identificar padrões e tendências que podem ser utilizados para o desenvolvimento do sistema de recomendação.
- **Desenvolvimento do sistema de recomendação:** Nesta etapa, serão desenvolvidos os algoritmos de recomendação. Os algoritmos serão avaliados em termos de precisão, cobertura e diversidade.
- **Implementação do sistema de recomendação:** Nesta etapa, o sistema de recomendação será implementado. O sistema será disponibilizado para usuários em um ambiente de teste.

2. Importando as bibliotecas

Importamos as bibliotecas: numpy, pandas, seaborn e sklearn. Com isso, é possível fazer uma análise de dados e representação dos dados e para a construção de um modelo de recomendação de livros. Essas ferramentas serão a base fundamental para explorar avaliações passadas e criar um sistema eficaz de recomendação.

3. Importando os Datasets

Importamos os arquivos contendo as informações dos filmes, as avaliações, os identificadores e as palavras-chave, e todas elas tendo um relacionamento, para conseguirmos conversar as bases e fazer a previsão com mais precisão.

4. Análise exploratória e preparação dos dados

Realizamos uma análise exploratória da base de dados, com a finalidade de detectar padrões e entender os dados que estamos trabalhando, a fim de facilitar na hora de preparar a base de dados para o modelo de Machine Learning, pois saberemos o que influencia em outras colunas, e também, colunas que poderíamos tirar, que afetaria positivamente na performance de processamento do modelo, e não tirando sua precisão.

5. Criação do sistema de recomendação

Criamos um DataFrame para realizar a matriz de avaliações dos usuários e filmes. Em seguida, criamos um pivot table para representar as avaliações dos usuários e filmes. Substituímos os valores ausentes por 0.

Utilizamos o algoritmo KNN para analisar os vizinhos mais próximos em busca de similaridade. O modelo foi treinado utilizando a métrica de similaridade de cosseno.

A partir do modelo treinado, criamos uma função de recomendação que utiliza as informações de sinopse e gênero dos filmes para afinar as recomendações. Utilizamos a biblioteca SentenceTransformer para calcular a similaridade de cosseno entre os filmes.

6. Exemplos de recomendação

Utilizamos a função de recomendação para sugerir filmes com base em filmes de referência selecionados, e com base em notas antigas que o usuário tem em nossa base de dados, usando um modelo de Machine Learning conhecido como KNN, que basicamente, detecta padrões, e com isso, “agrupa” coisas que são parecidas.

7. Conclusão

O projeto de recomendação de filmes utilizando Machine Learning demonstrou eficácia ao proporcionar sugestões personalizadas, enriquecendo a experiência do usuário e destacando o potencial impacto positivo da inteligência artificial na personalização de conteúdo cinematográfico.

8. Desafios

Calcular a acurácia no projeto revelou-se desafiador devido à natureza subjetiva das preferências de filmes, dificultando a obtenção de métricas precisas. Essa complexidade destaca a necessidade de aprimoramentos contínuos na avaliação de recomendações.

9. Conclusões

Sobre a pesquisa em conclusão, a pesquisa sobre recomendação de filmes por meio de Machine Learning evidencia não apenas os avanços alcançados na personalização de sugestões, mas também ressalta a constante necessidade de superar desafios na avaliação de desempenho. O caminho à frente requer inovação contínua para aprimorar a precisão das recomendações e, assim, proporcionar experiências mais enriquecedoras aos usuários.

10. Trabalhos futuros

Para trabalhos futuros, sugere-se explorar técnicas avançadas de aprendizado de máquina para lidar com preferências dinâmicas, e melhorar a maneira com que foi calculada a acurácia do modelo, para termos uma real noção de sua performance, e com isso, podemos aplicá-lo em situações reais.

11. Links úteis

- <https://youtu.be/tqD4BSKTcpk>
- [GabrielSicari/projeto3 at master \(github.com\)](https://github.com/GabrielSicari/projeto3)

12. Referências

<https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.NearestNeighbors.html>

<https://pypi.org/project/sentence-transformers/>

<https://medium.com/nlplanet/two-minutes-nlp-sentence-transformers-cheat-sheet-2e9865083e7a>

