CSE508 - Information Retrieval Project
Book Recommendation System

Shivanshu Kumar
2019476

Karan Prasad Gupta
2020439

Rishav
2020569

Sarthak Dixit
2020574

D Likhith
2019038

Harsh Kashyap
2020434

1. Introduction

We try to develop an innovative book recommendation system, providing users with a personalized and enriching reading experience.
Users face the following challenges:

- User Struggle: Overwhelmed by book choices, users face information overload.
- Engagement Boost: Personalized recommender systems enhance user engagement by consistently delivering tailored recommendations.
- Business Impact: Implementing recommender systems on large platforms contributes to increased user retention, satisfaction, and revenue generation.
- Competitive Advantage: Accurate and personalised recommender systems give platforms a competitive edge, simplifying user content discovery and aligning with their preferences.

2. Related Works

1. https://www.goodreads.com/

A one-stop site for book recommendations and reviews. Registered users can comment or discuss with others and get personalised recommendations based on their reading history.

Among other features are book clubs, challenges, quizzes and polls.

2. https://ieeexplore.ieee.org/document/9579647

Published by PV Devika et al. 2021, the research paper authors discover methods including machine learning techniques like the K-nearest neighbors, Pearson's R Correlation Coefficient, Cosine Similarneighboursollaborative filtering to be efficient in deciding the best books for the user based on the query that is provided as the input into their system.

3. https://cs.carleton.edu/cs_comps/1617/book_rec/final-results/paper.pdf

Proposed in 2016 by Jinny Cho et al., the project on book recommendation system works on the data collected from across the web (Amazon Books, GoogleReads, etc.) and makes use of the Machine Learning classifier models such as Naive Bayes Classifier and the Maximum Entropy Classifier for their Content based approach and the K-nearest neighbor and UV decomposition for their Collaborative filtering based approach for determination of appropriate books according to the user query.

4. https://arxiv.org/abs/2302.00653

Published in 2023 by Jessie Caridad et al., the research delves deep into the field of AI and proposes a state-of-the-art model for the determination of the books as demanded by the user by making use of the newly developed models called BRAIN-L (Base of Reasoning in Artificial Intelligence with Natural Language), wherein they combine certain ML techniques with the

Case-Based reasoning in AI, and using evaluation metrics like Cosine similarity, Jaccard's Index, Soft Cosine etc.

5. Book Recommendation System using Association Rule Mining & Collaborative Filtering

The research paper discusses several Collaborative filtering algorithms like the Jaccard Distance and Pearson' Coefficient along with a novel technique known as Association Mining.

3. Techniques/Algorithms

Since we are going to build a content-based Book Recommendation system using cosine similarity in order to decide which books are most appropriate to the book which the user searches, here come the similarity scores. It is a numerical value that ranges [0, 1], which helps to determine how much two things are similar to each other on a scale of zero to one. We can calculate the similarity score with the help of cosine similarity, which finds the similarity b/w the text details of a book that the user searches for and the rest of all books. So, similarity score is the measure of similarity between b/w given text details of two books. There are a few more collaborative filtering-based techniques/algorithms, such as knn, Jaccard's index, Soft cosine, etc., that we intend to integrate into our information retrieval model.

The novelty in the project:

Adapting a multi-modal approach to information retrieval of books as required by the user could include utilizing not just textual data but also images and audio to better understand user preferences. This could involve recommending books based on the book covers, author interviews, or book trailers. Another novel quality that we intend to apply to our project is the idea of interactive recommendations, where we allow users to provide feedback during the recommendation process through interactive interfaces such as quizzes, surveys or preference sliders, which can help refine recommendations in real-time.

4. Evaluation Methods

- Accuracy Metrics: Utilize metrics like MSE (Mean Squared Error) and RMSE (Root Mean Squared Error) to quantify prediction accuracy
- F-1 Score, Precision: Reliably suggests a wide variety of genres to the user's liking based on their input. Measure the F1-Score of recommended items compared to the actual user preferences
- Content-based filtering and Clustering: Custom curation and organization of related books to the user's preferences

5. Key Contributions:

We plan to enhance data quality by addressing missing values and outliers. Extract relevant features from book metadata and user preferences. Implement diverse recommendation algorithms, optimizing them for accuracy. Develop an intuitive user interface for easy interaction, and consider scalability for efficiently handling a large user base and item volume.