BOOK RECOMMENDATION SYSTEM USING HYBRID CONTENT AND COLLABORATIVE FILTERING TECHNIQUES

Group ID: CSE 25-72
Guide: Mr. Umang Rastogi

Team Members:
Vidyush Singh
Urvi Gupta
Tripti Singh

Course Outcomes (COs)

✅ CO1:

Understand the fundamentals of recommender systems and differentiate between content-based, collaborative, and hybrid filtering techniques.

✅ CO2:

Apply machine learning techniques like TF-IDF, cosine similarity, and matrix operations to build a content-based recommendation system.

✅ CO3:

Analyze and preprocess real-world datasets (Book-Crossing) for cleaning, feature extraction, and vectorization.

✅ CO4:

Implement and evaluate a hybrid recommendation model using performance metrics such as precision, recall, F1-score, and MAE.

✅ CO5:

Develop a robust, user-oriented book recommendation application that solves cold-start and sparsity issues using hybrid filtering.

✅ CO6:

Communicate and present technical solutions clearly through structured documentation and performance evaluations.

Program Outcomes (POs)

PO No.	Program Outcome	Mapping Explanation
PO1	Engineering Knowledge	Uses core CS/AI concepts like ML, vectorization, and similarity metrics.
PO2	Problem Analysis	Identifies cold-start, data sparsity, and accuracy issues in recommender systems.
PO3	Design/Development of Solutions	Designs and implements a hybrid filtering solution with measurable improvements.
PO4	Use of Modern Tools	Uses Scikit-learn, TF-IDF, Python, Surprise library for model development.
PO5	Modern Tool Usage	Demonstrates data analysis, modeling, and deployment techniques.
PO8	Ethics	(If discussed) In real-world apps, ethical use of user data could be mentioned.
PO9	Teamwork	Team project with defined roles and contributions.
PO10	Communication	Clear documentation, code organization, and explanation of results.
PO11	Project Management	Manages dataset selection, model building, evaluation, and presentation.

PROJECT ABSTRACTION

- Aim: To develop a personalized book recommendation system.
- Uses hybrid filtering combining collaborative and content-based filtering.
- Enhances reading experience through tailored suggestions.
- Addresses cold-start problems and improves user engagement.

OVERVIEW

- Personalized recommendations based on user preferences and book features.
- Integrates collaborative filtering (CF) and content-based filtering (CBF).
- Handles data sparsity and cold-start issues effectively.
- Ensures diversity, scalability, and improved accuracy.

KEY OBJECTIVES

- Build a hybrid recommendation engine.
- Use Book-Crossing dataset for training and evaluation.
- Apply machine learning for user behavior analysis.
- Enhance precision, recall, and user satisfaction.

APPROACH

- CF: Analyzes user-item interaction history.
- CBF: Extracts book attributes using TF-IDF.
- Hybrid: Merges CF and CBF outputs.
- Model trained and evaluated using standard metrics.

METHODOLOGY

- Dataset: Book-Crossing Dataset (278k users, 271k books).
- Preprocessing: Cleaning, normalization, TF-IDF encoding.
- Filtering Models: CF, CBF, Hybrid.
- Tools: Python, Scikit-learn, Surprise library.

MODEL DESIGN

- CF Architecture: User-based and Item-based CF.
- CBF Architecture: Feature vectors of book metadata.
- Hybrid Architecture: Combines both outputs for final suggestions.
- Similarity measures: Cosine, Pearson correlation.

EXPECTED OUTCOMES

- High accuracy and personalized book suggestions.
- Hybrid model achieves:
  - Precision: 0.85
  - Recall: 0.88
  - F1-Score: 0.86
  - MAE: 0.75
  - Real-time recommendation capabilities.

TESTING AND REFINEMENT

- 80:20 train-test data split.
- Evaluation Metrics: Precision, Recall, F1, MAE.
- Hybrid model consistently outperforms CF and CBF.
- Refinement based on feedback and performance metrics.

SUSTAINABLE DEVELOPMENT GOALS

- SDG 4: Quality Education - Personalized learning.
- SDG 9: Innovation - AI-based recommendation system.
- SDG 10: Reduced Inequality - Equal access to relevant content.

CONCLUSION

- Hybrid filtering is robust and accurate.
- Combines CF's community insight with CBF's feature analysis.
- Future Scope:
  - Deep learning models (e.g., BERT).
  - Real-time updates and user feedback integration.
  - Cross-domain recommendations.
  - Explainable AI for user trust.

THANK YOU