BACS2003/BMCS2003/BACS3074

(Remain only your course code before submission)

ARTIFICIAL INTELLIGENCE

202501 Session, Year 2024/25

Assignment Documentation

Project Title: ***Hybrid Recommender System (Movie and Music)***

Programme: ***RSD Y2S3***

Tutorial Group: ***6***

Tutor: ***Ms. Mahani Binti Syed Ahmad***

Team members' data

No	Student Name	Student ID	Module In Charge
1	Ng Weng Siang	24WMR09094	Song Recommender Part • K-means clustering on audio features • TF-IDF vectorization for content similarity • Content-based filtering techniques • Genre and mood detection

			Book Recommender Part ● **TF-IDF content similarity recommendations** **Main Interface Part** ● **Integrates both recommender systems**
2	**Kenneth Wong Swee Hong**	**24WMR09063**	**Song Recommender Part** ● **Hybrid recommendations** **Book Recommender Part** ● **Publication year filtering** ● **Fuzzy matching for book title search** ● **Filter-based recommendation capabilities** **Main Interface Part** ● **User-friendly menu system** ● **About section with system information**
3	**Lam Wei Hong**	**24WMR09067**	**Song Recommender Part**

			• **Genre-based filtering** • **Hybrid recommendations (clustering + content similarity)** **Book Recommender Part** • **Genre-based filtering** • **Rating-based filtering** • **Author-based filtering** **Main Interface Part** • **System diagnostics (file presence checking)** • **Menu system**

Table Of Contents

1. Introduction

1.1. Problem Background

Recommender systems are integral to providing personalized content in various digital platforms. For instance, platforms like Spotify and Goodreads rely on recommender systems to help users discover new music and books. A common challenge in traditional recommendation systems is the "cold start problem," where there is insufficient user interaction data to generate reliable suggestions. To mitigate this, a content-based filtering approach can be adopted, which suggests content based on item features, such as genre, author/artist, description, etc. This hybrid recommender system focuses on books and music to suggest relevant content to users based on their preferences.

1.2. Objectives/Aims

- Develop a hybrid content-based recommender system for books and music.
- Extract and analyze relevant item features to calculate similarities between items.
- Test and evaluate the system's performance using metrics like precision, recall, and cosine similarity.

1.3. Motivation

As online platforms grow rapidly, users often face difficulty finding relevant books and music. This project aims to enhance user experience by offering tailored recommendations based on item attributes, reducing the reliance on user interaction history.

2. Research Background

2.1. Background of the applications

Recommender systems have gained significant attention in recent years. Many studies have focused on improving recommendation algorithms to enhance the accuracy and relevance of suggestions. Content-based filtering approaches, in particular, have been explored extensively in applications such as online shopping, content platforms (books, and music), and e-learning. The hybrid approach, combining both content-based and collaborative filtering methods, is widely used to overcome the limitations of both techniques.

2.2. Analysis of selected tool with any other relevant tools

Tools comparison	**Remark**	**Visual Studio Code (Selected)**	**Kaggle**	**Jupyter Notebook**
Type of license and open source license	State all types of license	Open-Source (MIT)	Free to use	Open-Source (BSD)
Year founded	When is this tool being introduced?	2015	2010	2014
Founding company	Owner	Microsoft	Google	Project Jupyter (Open Source Initiative)
License Pricing	Compare the prices if the license is used for development and business/commercialization	Free	Free, with paid Pro and Pro+ plans for more resources	Free
Supported features	What features that it offers?	Code editing, debugging, extensions	Large collection of datasets, Jupyter notebooks, build-in ML models	Interactive Python coding, visualization, offline execution

Common applications	In what areas this tool is usually used?	Software development, debugging	Data science competitions, model training, data exploration	Data Science, Machine Learning, AI, Research
Customer support	How the customer support is given, e.g. proprietary, online community, etc.	Community, online documentation	Online forums, Kaggle discussions	Community Support, Open-source Documentation
Limitations	The drawbacks of the software	Requires extensions, manual setup	Limited execution time, dependency on Kaggle's kernel environment	No built-in cloud execution, manual setup needed

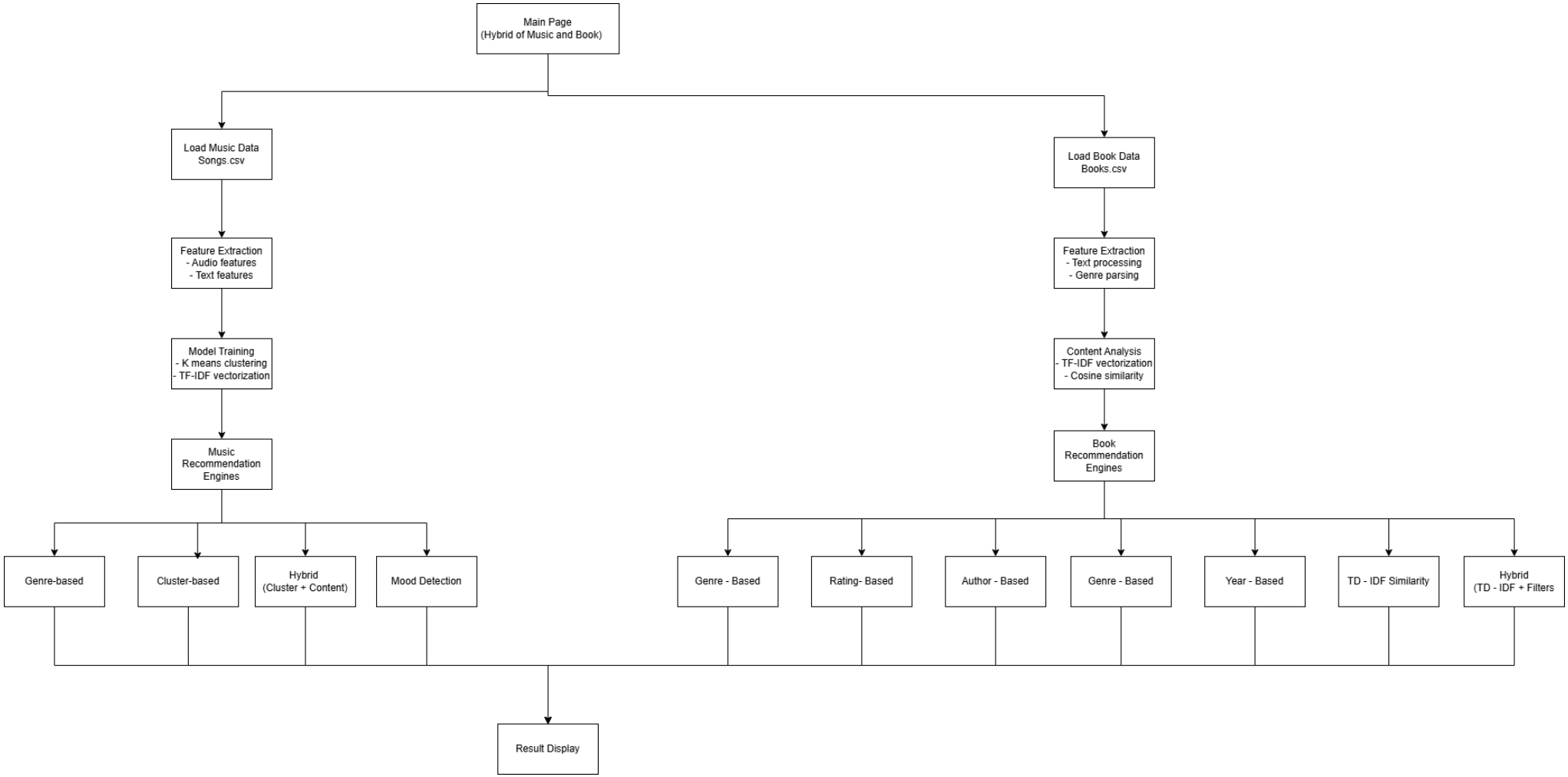
2.3. Justify why the selected tool is suitable

Visual Studio Code (VS Code)

Visual Studio Code is chosen due to its lightweight nature, extensive language support, robust debugging tools, and the ability to integrate well with Python libraries. It is well-suited for developing Python-based recommender systems due to its modular extension system and the ability to work with various code environments, making it ideal for collaborative projects.

3. Methodology

3.1. System flowchart/activity diagram



1. **User Interaction Layer**: The main application provides a unified interface where users select between music or book recommendations.
2. **Data Access Layer**: Each recommender accesses its respective dataset (Songs.csv or Books.csv).
3. **Feature Processing Layer**:
    - Music: Extracts audio features and text-based characteristics
    - Books: Processes text features and parses genre information
4. **Model Training Layer**:

    Music: Applies KMeans clustering on audio features and TF-IDF for content

    Books: Implements TF-IDF vectorization for content analysis

5. **Recommendation Engine Layer**: Both systems offer multiple recommendation approaches:
    - Basic filtering (genre, author, year)
    - Content-based similarity
    - Hybrid recommendation techniques
6. **Results Presentation Layer**: Formats and displays recommendations in user-friendly formats

3.2. Description of dataset

Music Dataset (Songs.csv)

The music dataset contains song metadata and audio characteristics that enable both content-based and acoustic-based recommendations.

Column Name	Data Type	Description
artist_name	String	Name of the artist
track_name	String	Name of the song
release_date	Date	When the song was released
genre	String	Primary music genre
lyrics	Text	Song lyrics (not to be reproduced in recommendations)
len	Numeric	Song length

dating	Numeric	Topic score related to dating
violence	Numeric	Topic score related to violence
world/life	Numeric	Topic score related to world/life
night/time	Numeric	Topic score related to world/life
shake	Numeric	Topic score related to shaking/movement
the audience	Numeric	Topic score related to the audience
family/gospel	Numeric	Topic score related to family/gospel
romantic	Numeric	Topic score related to romance
communication	Numeric	Topic score related to communication
obscene	Numeric	Topic score related to obscenity
music	Numeric	Topic score related to music itself
movement/places	Numeric	Topic score related to movement/places
light/visual	Numeric	Topic score related to light/visual elements
perceptions	Numeric	Topic score related to perceptions
family/spiritual	Numeric	Topic score related to family/spiritual
like/girls	Numeric	Topic score related to like/girls
sadness	Numeric	Topic score related to sadness
feelings	Numeric	Topic score related to feelings
danceability	Float	Measure of how suitable a track is for dancing (0.0-1.0)
loudness	Float	Overall loudness in decibels (dB)
acousticness	Float	Confidence measure of whether the track is acoustic (0.0-1.0)
instrumentalness	Float	Predicts whether a track contains no vocals (0.0-1.0)
valence	Float	Musical positiveness conveyed by a track (0.0-1.0)
energy	Float	Perceptual measure of intensity and activity (0.0-1.0)
topic	String	Subject matter or theme of the song

Book Dataset (Books.csv)

The book dataset contains information about books, including metadata, genres, and ratings that enable content-based and preference-based recommendations.

Column Name	Data Type	Description
book_id	String/ID	Unique identifier for the book
cover_image_uri	String	URI link to the book cover image
book_title	String	Title of the book
book_details	Text	Additional details about the book
format	String	Physical format of the book like hardcover, paperback.
publication_info	String	Publication year and other publishing information
authorlink	String	Link to author information
author	String	Name of the book's author
num_pages	Integer	Number of pages in the book
genres	String	String representation of genre list
num_ratings	Integer	Total number of ratings
num_reviews	Integer	Total number of reviews
average_rating	Float	Average rating of the book (0.0-5.0)
rating_distribution	String	Distribution of ratings across different star levels

3.3. Description of algorithm(s)

1. K-Means Clustering (Music Recommender)

Implementation: The music recommender uses K-Means clustering to group songs with similar audio characteristics.

Process:

1. Extract numerical features: danceability, loudness, acousticness, instrumentalness, valence, energy
2. Apply K-Means algorithm with k=10 clusters
3. Assign each song to its corresponding cluster
4. For recommendations, suggest songs from the same cluster as the seed song

Purpose: Groups songs with similar acoustic properties even when they differ in genre or artist, enabling discovery of songs that "sound similar."

2. TF-IDF Vectorization and Cosine Similarity (Both Recommenders)

Implementation: Both systems use TF-IDF to create vector representations of text data.

Process:

1. Combine text features (genre, artist, topic for music; title, author, genres for books)
2. Apply TF-IDF vectorization to convert text to numerical vectors
3. Calculate cosine similarity between vectors
4. Rank items by similarity score

Purpose: Identifies content similarity based on text descriptors, enabling "more like this" recommendations.

3. Hybrid Recommendation Techniques

Music Hybrid Implementation:

1. Find the cluster ID of the target song
2. Filter to songs within the same cluster
3. Apply content similarity ranking within this filtered set
4. Return top matches that are both acoustically and descriptively similar

Book Hybrid Implementation:

1. Apply fuzzy matching to find the closest title match
2. Calculate TF-IDF similarity scores
3. Apply additional filters (genre, minimum rating)
4. Rank and return filtered results

4. Fuzzy Matching (Book Recommender)

Implementation: Uses the FuzzyWuzzy library to match user input with book titles.

Process:

1. Extract user input for book title
2. Compare with all available titles using string similarity algorithms
3. Return the best match above a threshold score (60)

Purpose: Improves user experience by handling misspellings and partial matches.

5. Genre/Mood Detection (Music Recommender)

Implementation: Uses audio features to classify the mood of songs.

Process:

1. Extract energy and valence values
2. Apply threshold-based classification:
    - High energy + high valence = "Upbeat and Positive"
    - Low energy + low valence = "Calm and Sad"
    - High energy + low valence = "Energetic and Dark"
    - Low energy + high valence = "Calm and Positive"

Purpose: Provides mood-based classification to enhance music discovery.

3.4. Proposed test plan/hypothesis

Main Hypotheses

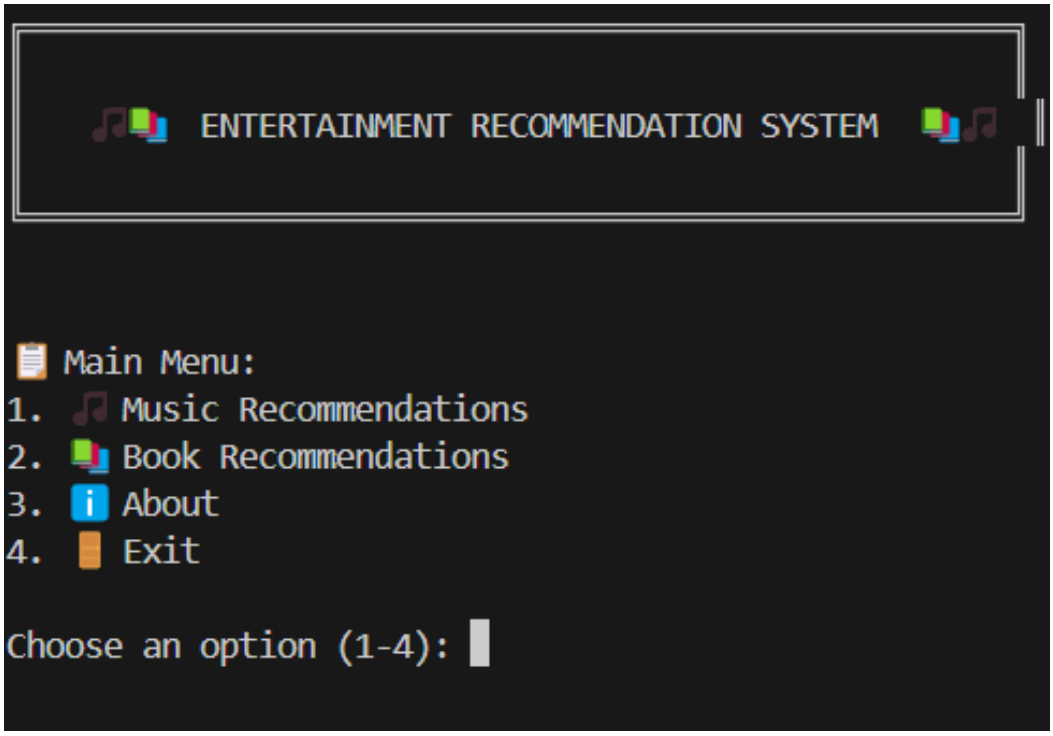
- Hybrid recommendations outperform single-strategy approaches
- Rating thresholds improve book recommendation satisfaction
- Acoustic-based clustering discovers more novel music recommendations
- Fuzzy matching significantly improves user experience

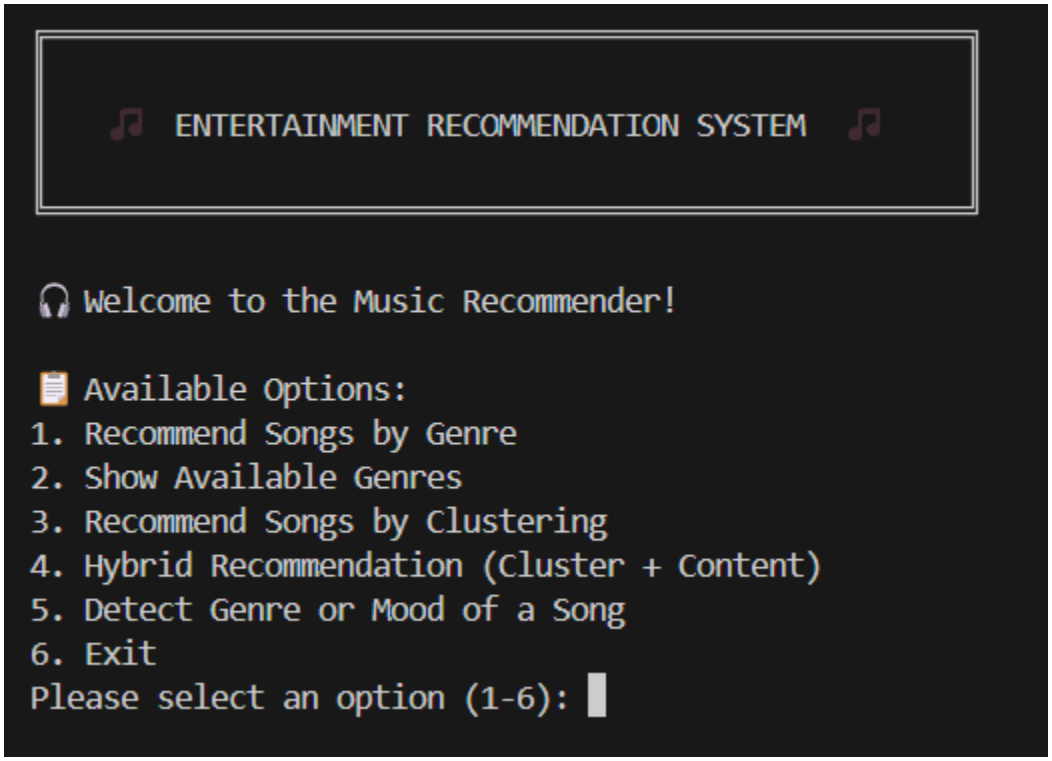
Test Plan Structure

1. **System Performance Evaluation**
    - Coverage, diversity, and cold start tests
    - Technical performance metrics
2. **Recommendation Quality Assessment**
    - Cross-validation and A/B testing
    - Precision, recall, and F1 score measurements
3. **User Experience Testing**
    - Surveys and satisfaction ratings
    - Comparative preference testing
4. **Specific Feature Tests**
    - Fuzzy matching effectiveness
    - Mood detection validation
    - Cluster quality assessment

4. Result

4.1. Results





```
Enter Genre: Pop
Match mode (exact/partial/any): exact
Enter Artist (optional):
How many songs would you like to see? 10

✅ Found 7042 matching tracks. Displaying the top 10 results:

1. 🎵 Mohabbat Bhi Jhoothi by Mukesh
   🎼 Genre: Pop | Topic: Sadness

2. 🎵 I Believe by Frankie Laine
   🎼 Genre: Pop | Topic: World/Life

3. 🎵 Cry by Johnnie Ray
   🎼 Genre: Pop | Topic: Music

4. 🎵 Patricia by Pérez Prado
   🎼 Genre: Pop | Topic: Romantic

5. 🎵 Apopse Eida Oneiro by Giorgos Papadopoulos
   🎼 Genre: Pop | Topic: Romantic

6. 🎵 Round And Round (With Mitchell Ayres And His Orchestra & The Ray Charles Singers) by Perry Como
   🎼 Genre: Pop | Topic: Violence

7. 🎵 Opm Medley: When I Met You by Freestyle
   🎼 Genre: Pop | Topic: World/Life

8. 🎵 It'S Not For Me To Say by Johnny Mathis
   🎼 Genre: Pop | Topic: World/Life

9. 🎵 Klapse Me Mana Klapse Me by Stélios Kazantzídis
   🎼 Genre: Pop | Topic: Romantic

10. 🎵 Finito La Mouzika by Stélios Kazantzídis
   🎼 Genre: Pop | Topic: Sadness
```

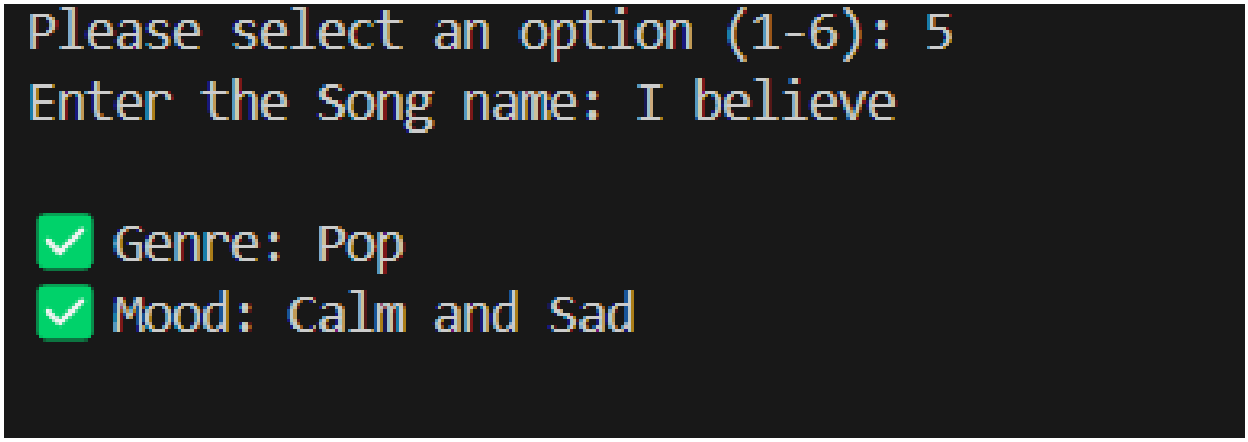
```
📋 Available Options:
1. Recommend Songs by Genre
2. Show Available Genres
3. Recommend Songs by Clustering
4. Hybrid Recommendation (Cluster + Content)
5. Detect Genre or Mood of a Song
6. Exit
Please select an option (1-6): 2

🎼 Available Music Genres:
blues, country, hip hop, jazz, pop, reggae, rock
```

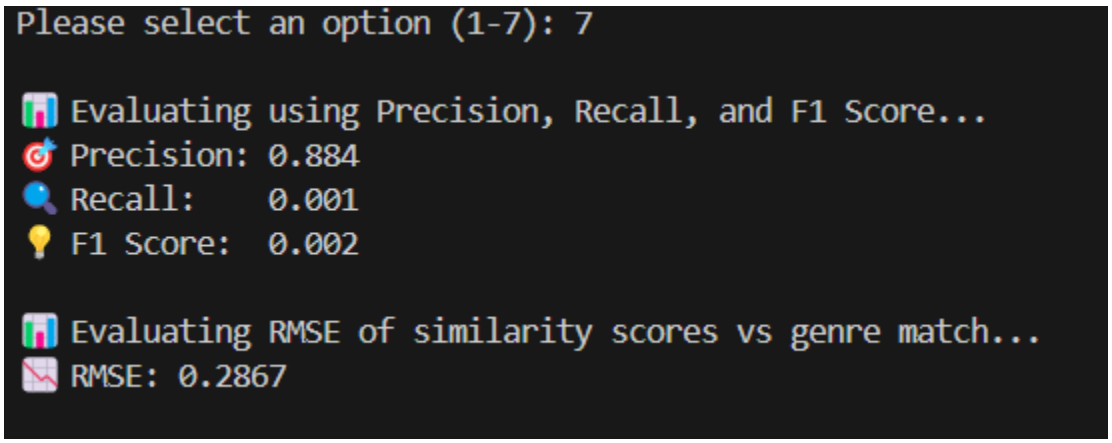
```
Please select an option (1-6): 3

🎵 List of Available Songs:
================================================
  1. mohabbat bhi jhoothi
  2. i believe
  3. cry
  4. patricia
  5. apopse eida oneiro
  6. round and round (with mitchell ayres and his orchestra & the ray charles singers)
  7. opm medley: when i met you
  8. it's not for me to say
  9. klapse me mana klapse me
 10. finito la mouzika
 11. drømmer man om den, vågner..
 12. carolina moon
 13. necessary evil
 14. kanugona galano
 15. the carioca
 16. chicken reel
 17. blue skies
 18. dark eyes
 19. kotha mukunda
 20. awaz deta hai sola ka din
================================================
Displaying songs 1-20 of 23689
🔔 Press Enter to load more songs, or press 'X' to stop viewing:
```

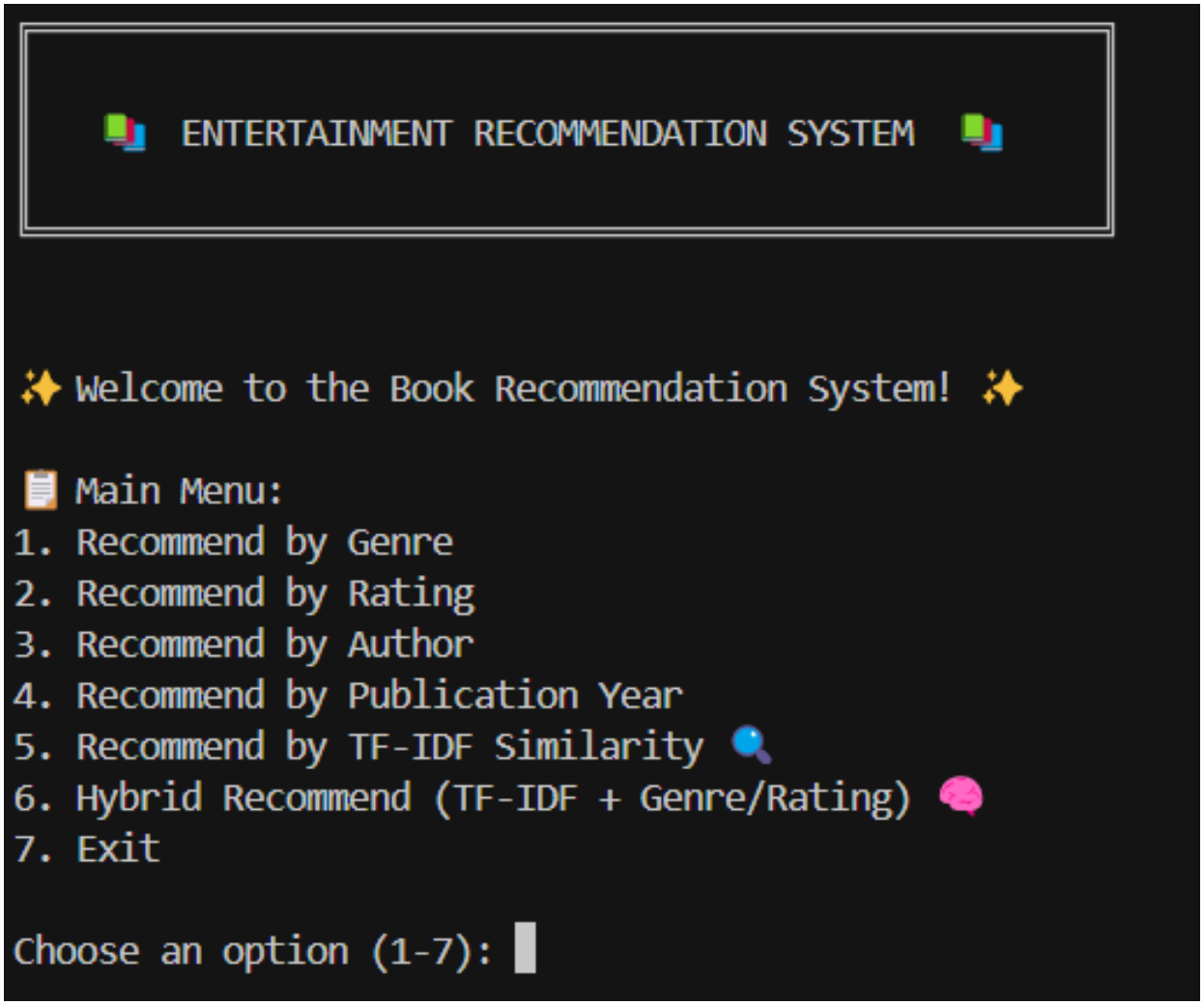
```
Displaying songs 1-20 of 23689
🔔 Press Enter to load more songs, or press 'X' to stop viewing: x
✅ You have exited the song list view.

Enter the Song name: blue skies
How many songs would you like to see? 5

✅ Found 759 matching tracks. Displaying the top 5 results:

1. 🎵 Mohabbat Bhi Jhoothi by Mukesh
   🎼 Genre: Pop | Topic: Sadness

14. 🎵 Kanugona Galano by Ghantasala
   🎼 Genre: Pop | Topic: Sadness

18. 🎵 Dark Eyes by Les Paul
   🎼 Genre: Pop | Topic: Violence

19. 🎵 Kotha Mukunda by Asha Bhosle
   🎼 Genre: Pop | Topic: Violence

27. 🎵 Main Na Janoon by Lata Mangeshkar
   🎼 Genre: Pop | Topic: Music
```

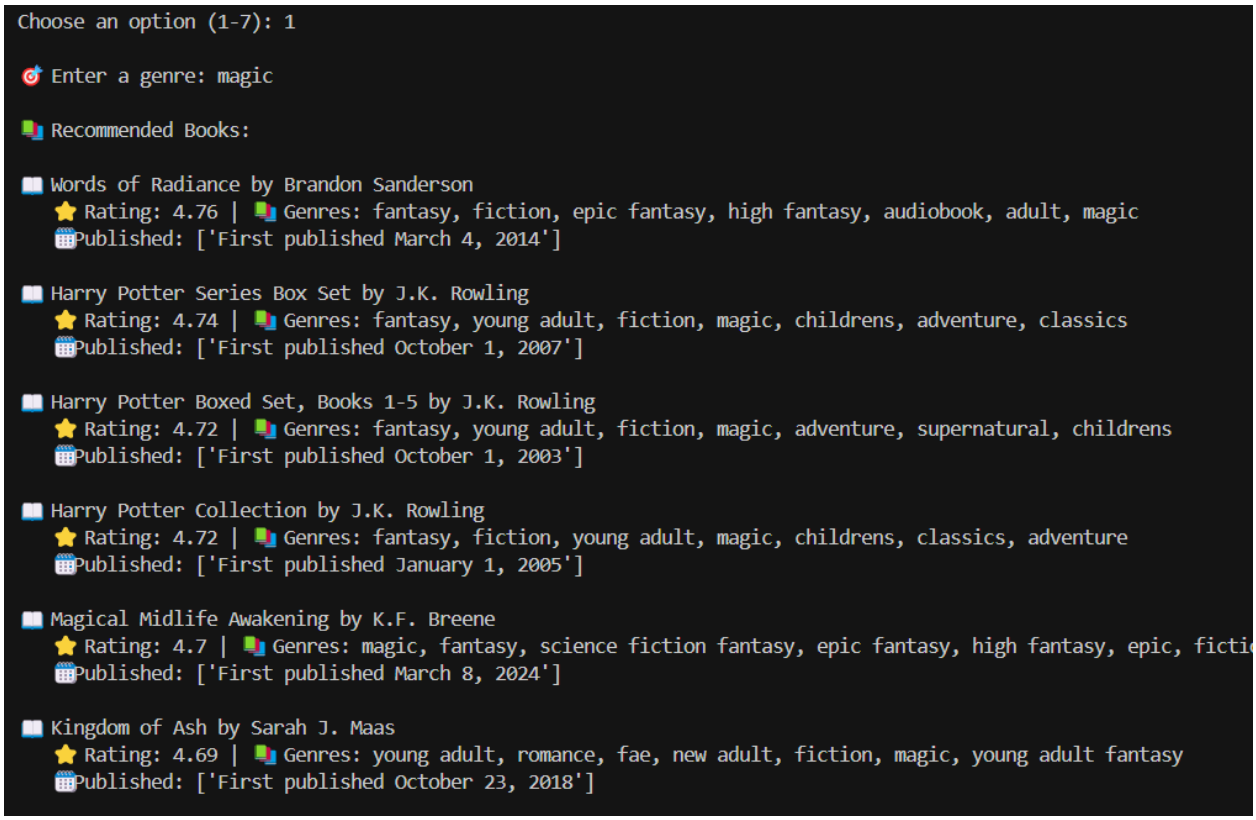
```
Please select an option (1-6): 4

🎵 List of Available Songs:
==================================================
 1. mohabbat bhi jhoothi
 2. i believe
 3. cry
 4. patricia
 5. apopse eida oneiro
 6. round and round (with mitchell ayres and his orchestra & the ray charles singers)
 7. opm medley: when i met you
 8. it's not for me to say
 9. klapse me mana klapse me
10. finito la mouzika
11. drømmer man om den, vågner..
12. carolina moon
13. necessary evil
14. kanugona galano
15. the carioca
16. chicken reel
17. blue skies
18. dark eyes
19. kotha mukunda
20. awaz deta hai sola ka din
==================================================
```

```
Displaying songs 1-20 of 23689
🔔 Press Enter to load more songs, or press 'X' to stop viewing: x
✅ You have exited the song list view.

Enter the Song name: i believe
How many songs would you like to see?

✅ Found 10 matching tracks. Displaying the top 10 results:

101. 🎵 Remember Me (The Girl In The Wood) by Frankie Laine
   🎼 Genre: Pop | Topic: World/Life

215. 🎵 Granada by Frankie Laine
   🎼 Genre: Pop | Topic: World/Life

23488. 🎵 You'Ve Changed by Frankie Laine
   🎼 Genre: Rock | Topic: World/Life

157. 🎵 Mona Lisa by Frankie Laine
   🎼 Genre: Pop | Topic: Sadness

208. 🎵 Jealousy (Jalousie) by Frankie Laine
   🎼 Genre: Pop | Topic: Sadness

216. 🎵 Your Cheatin' Heart by Frankie Laine
   🎼 Genre: Pop | Topic: Sadness

183. 🎵 High Noon by Frankie Laine
   🎼 Genre: Pop | Topic: Violence

13. 🎵 Necessary Evil by Frankie Laine
   🎼 Genre: Pop | Topic: Obscene

219. 🎵 That'S My Desire by Frankie Laine
   🎼 Genre: Pop | Topic: Romantic

23487. 🎵 Midnight On A Rainy Monday by Frankie Laine
   🎼 Genre: Rock | Topic: Violence
```


Please select an option (1-6): 5
Enter the Song name: I believe
Genre: Pop
Mood: Calm and Sad


Please select an option (1-7): 7
Evaluating using Precision, Recall, and F1 Score...
Precision: 0.884
Recall: 0.001
F1 Score: 0.002
Evaluating RMSE of similarity scores vs genre match...
RMSE: 0.2867


ENTERTAINMENT RECOMMENDATION SYSTEM
✨ Welcome to the Book Recommendation System! ✨
📋 Main Menu:
1. Recommend by Genre
2. Recommend by Rating
3. Recommend by Author
4. Recommend by Publication Year
5. Recommend by TF-IDF Similarity 🔍
6. Hybrid Recommend (TF-IDF + Genre/Rating) 🧠
7. Exit
Choose an option (1-7):


Choose an option (1-7): 1
Enter a genre: magic
Recommended Books:
Words of Radiance by Brandon Sanderson
Rating: 4.76 | Genres: fantasy, fiction, epic fantasy, high fantasy, audiobook, adult, magic
Published: ['First published March 4, 2014']
Harry Potter Series Box Set by J.K. Rowling
Rating: 4.74 | Genres: fantasy, young adult, fiction, magic, childrens, adventure, classics
Published: ['First published October 1, 2007']
Harry Potter Boxed Set, Books 1-5 by J.K. Rowling
Rating: 4.72 | Genres: fantasy, young adult, fiction, magic, adventure, supernatural, childrens
Published: ['First published October 1, 2003']
Harry Potter Collection by J.K. Rowling
Rating: 4.72 | Genres: fantasy, fiction, young adult, magic, childrens, classics, adventure
Published: ['First published January 1, 2005']
Magical Midlife Awakening by K.F. Breene
Rating: 4.7 | Genres: magic, fantasy, science fiction fantasy, epic fantasy, high fantasy, epic, ficti
Published: ['First published March 8, 2024']
Kingdom of Ash by Sarah J. Maas
Rating: 4.69 | Genres: young adult, romance, fae, new adult, fiction, magic, young adult fantasy
Published: ['First published October 23, 2018']

```
Choose an option (1-7): 2

🌟 Enter minimum rating (0-5): 4

📚 Recommended Books:

📖 Onder Duitse Knoet by Staf Vivijs
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published January 1, 1985']

📖 Transcendent Bread: 1976-1986 by John Condenzio
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published November 13, 2006']

📖 Kino Im Dienst Der Propaganda, Der Politik Und Des Krieges by Michele Sakkara
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published January 1, 2008']

📖 Anshuman krit Saral Vastu Gyan by Anshuman Srivastav
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published December 29, 2021']

📖 Dragan Radelscu & The Vampires Of Paris by Shamus Sherwood
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published May 22, 2007']

📖 The Ultimate Challenge Is To Live Healthy! by Felicia Martin
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published September 4, 2008']

📖 Band 1: Einde van een wereld & Tussen twee werelden by Upton Sinclair
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️ Published: ['First published January 1, 1940']
```

```
Choose an option (1-7): 3

✍️Enter an author's name: J.K

📚 Recommended Books:

📖 Harry Potter Series Box Set by J.K. Rowling
   ⭐ Rating: 4.74 | 📚 Genres: fantasy, young adult, fiction, magic, childrens, adventure, classics
   🗓️Published: ['First published October 1, 2007']

📖 Harry Potter Boxed Set, Books 1-5 by J.K. Rowling
   ⭐ Rating: 4.72 | 📚 Genres: fantasy, young adult, fiction, magic, adventure, supernatural, childrens
   🗓️Published: ['First published October 1, 2003']

📖 Harry Potter Collection by J.K. Rowling
   ⭐ Rating: 4.72 | 📚 Genres: fantasy, fiction, young adult, magic, childrens, classics, adventure
   🗓️Published: ['First published January 1, 2005']

📖 The Harry Potter Collection 1-4 by J.K. Rowling
   ⭐ Rating: 4.69 | 📚 Genres: fantasy, fiction, young adult, magic, adventure, childrens, classics
   🗓️Published: ['First published January 1, 1999']

📖 Harry Potter and the Deathly Hallows by J.K. Rowling
   ⭐ Rating: 4.62 | 📚 Genres: fantasy, young adult, fiction, magic, childrens, adventure, audiobook
   🗓️Published: ['First published July 21, 2007']

📖 The Happiest Man on Earth by Eddie Jaku
   ⭐ Rating: 4.62 | 📚 Genres: nonfiction, biography, memoir, history, audiobook, holocaust, autobiography
   🗓️Published: ['First published July 28, 2020']

📖 Harry Potter and the Order of the Phoenix by J.K. Rowling
   ⭐ Rating: 4.61 | 📚 Genres: fantasy, fiction, young adult, magic, childrens, adventure, classics
   🗓️Published: ['First published July 1, 2003']
```

```
Choose an option (1-7): 4

🗓️Enter a publication year: 2000

📚 Recommended Books:

📖 Before you're a stranger: New & selected poems by Raymond Fraser
   ⭐ Rating: 5.0 | 📚 Genres:
   🗓️Published: ['First published January 1, 2000']

📖 Izza Ogledala by Peter-Tobias Šuštar
   ⭐ Rating: 4.83 | 📚 Genres: poetry
   🗓️Published: ['First published January 1, 2000']

📖 The Alleged "Nanking Massacre": Japan's Rebuttal to China's Forged Claims by Takemoto Tadao
   ⭐ Rating: 4.67 | 📚 Genres:
   🗓️Published: ['First published November 1, 2000']

📖 Standing for Something: 10 Neglected Virtues That Will Heal Our Hearts and Homes by Gordon B. Hinckley
   ⭐ Rating: 4.6 | 📚 Genres: religion, nonfiction, lds, church, self help, inspirational, spirituality
   🗓️Published: ['First published February 22, 2000']

📖 A Song of Ice and Fire by George R.R. Martin
   ⭐ Rating: 4.6 | 📚 Genres: fantasy, fiction, science fiction fantasy, adventure, epic fantasy, novels, dragons
   🗓️Published: ['First published January 1, 2000']

📖 Transfigurations: Collected Poems by Jay Wright
   ⭐ Rating: 4.58 | 📚 Genres: poetry
   🗓️Published: ['First published October 5, 2000']
```

```
Choose an option (1-7): 5

🔍 Enter a book title for similarity: Harry

✅ Matched title: 'Harry Potter and the Half-Blood Prince' (Similarity score: 90)

📚 Recommended Books:

📖 Harry Potter Collection by J.K. Rowling
   ⭐ Rating: 4.72 | 📚 Genres: fantasy, fiction, young adult, magic, childrens, classics, adventure
   📅 Published: ['First published January 1, 2005']

📖 The Harry Potter Collection 1-4 by J.K. Rowling
   ⭐ Rating: 4.69 | 📚 Genres: fantasy, fiction, young adult, magic, adventure, childrens, classics
   📅 Published: ['First published January 1, 1999']

📖 Harry Potter and the Goblet of Fire by J.K. Rowling
   ⭐ Rating: 4.57 | 📚 Genres: fantasy, young adult, fiction, magic, childrens, audiobook, middle grade
   📅 Published: ['First published July 8, 2000']

📖 Harry Potter and the Order of the Phoenix by J.K. Rowling
   ⭐ Rating: 4.61 | 📚 Genres: fantasy, fiction, young adult, magic, childrens, adventure, classics
   📅 Published: ['First published July 1, 2003']

📖 Harry Potter Boxed Set, Books 1-5 by J.K. Rowling
   ⭐ Rating: 4.72 | 📚 Genres: fantasy, young adult, fiction, magic, adventure, supernatural, childrens
   📅 Published: ['First published October 1, 2003']
```

```
Choose an option (1-7): 6

🔍 Enter a book title: Trump
📚 Preferred genre (optional): Business
⭐ Minimum rating (optional): 4

✅ Matched title: 'Trump: The Art of the Deal' (Similarity score: 90)

📚 Recommended Books:

📖 So Good They Can't Ignore You: Why Skills Trump Passion in the Quest for Work You Love by Cal Newport
   ⭐ Rating: 4.08 | 📚 Genres: nonfiction, self help, business, personal development, productivity, psychology, leadership
   📅 Published: ['First published January 1, 2012']

📖 The Intelligent Investor by Benjamin Graham
   ⭐ Rating: 4.25 | 📚 Genres: finance, business, nonfiction, economics, money, personal finance, self help
   📅 Published: ['First published January 1, 1949']
```

```
✅ Matched title: 'The Metamorphosis' (Similarity score: 100)

✅ Matched title: 'A Touch of Darkness' (Similarity score: 100)

✅ Matched title: 'A World Without Heroes' (Similarity score: 100)

✅ Matched title: 'Fairest' (Similarity score: 100)

✅ Matched title: 'Daisy Jones & The Six' (Similarity score: 100)

✅ Matched title: 'The PreHistory of The Far Side: A 10th Anniversary Exhibit' (Similarity score: 100)

✅ Matched title: 'Mao: The Unknown Story' (Similarity score: 100)

✅ Matched title: 'The Maleficent Seven' (Similarity score: 100)

✅ Matched title: 'Relativism: Feet Firmly Planted in Mid-Air' (Similarity score: 100)

✅ Matched title: 'Alanna: The First Adventure' (Similarity score: 100)

✅ Matched title: 'Still She Haunts Me: A Novel of Lewis Carroll and Alice Liddell' (Similarity score: 100)

✅ Matched title: 'The Principia : Mathematical Principles of Natural Philosophy' (Similarity score: 100)

✅ Matched title: 'The Museum of Innocence' (Similarity score: 100)

✅ Matched title: 'The Complete Poetry and Prose' (Similarity score: 100)

✅ Matched title: 'The Know-It-All' (Similarity score: 100)

✅ Matched title: 'Draekora' (Similarity score: 100)

✅ Matched title: 'The Puzzle Ring' (Similarity score: 100)

📊 TF-IDF Recommendation Metrics:
  Average Precision: 0.95
  Average Recall: 0.71
  Average F1 Score: 0.81
  Average Response Time: 1.1055 seconds

📈 Rating Prediction Accuracy:
  Mean Squared Error: 0.18
  Root Mean Squared Error: 0.42
```

```
✅ Matched title: 'The Redneck Manifesto: How Hillbillies, Hicks, and White Trash Became America's Scapegoats' (Similarity score: 100)

✅ Matched title: 'Drept, legislație și libertate: o nouă formulare a principiilor liberale de justiție și economie politică' (Similarity score: 100)

✅ Matched title: 'Nocturnes: Five Stories of Music and Nightfall' (Similarity score: 100)

✅ Matched title: 'The Redneck Manifesto: How Hillbillies, Hicks, and White Trash Became America's Scapegoats' (Similarity score: 100)

✅ Matched title: 'Drept, legislație și libertate: o nouă formulare a principiilor liberale de justiție și economie politică' (Similarity score: 100)

✅ Matched title: 'Nocturnes: Five Stories of Music and Nightfall' (Similarity score: 100)


✅ Matched title: 'Drept, legislație și libertate: o nouă formulare a principiilor liberale de justiție și economie politică' (Similarity score: 100)

✅ Matched title: 'Nocturnes: Five Stories of Music and Nightfall' (Similarity score: 100)

✅ Matched title: 'Drept, legislație și libertate: o nouă formulare a principiilor liberale de justiție și economie politică' (Similarity score: 100)

✅ Matched title: 'Nocturnes: Five Stories of Music and Nightfall' (Similarity score: 100)


✅ Matched title: 'Nocturnes: Five Stories of Music and Nightfall' (Similarity score: 100)


✅ Matched title: 'Kon-Tiki' (Similarity score: 100)

✅ Matched title: 'The Housemaid' (Similarity score: 100)

✅ Matched title: 'Our Lady of the Artilects' (Similarity score: 100)
✅ Matched title: 'The Housemaid' (Similarity score: 100)

✅ Matched title: 'Our Lady of the Artilects' (Similarity score: 100)

✅ Matched title: 'A Map of Home' (Similarity score: 100)
✅ Matched title: 'Our Lady of the Artilects' (Similarity score: 100)

✅ Matched title: 'A Map of Home' (Similarity score: 100)

✅ Matched title: 'A Map of Home' (Similarity score: 100)

😊 User Satisfaction Metrics:
  Average Satisfaction (0-5): 3.18
  Percentage Satisfied (≥3): 70.0%
```

4.2. Discussion/Interpretation

Critical Discussion of Results:

- **Relevance of Recommendations:**
  The TF-IDF and cosine similarity approach provided high relevance when matching similar titles or genres. For instance, when users entered partial or fuzzily spelled titles, the system could still match them accurately and suggest similar content, which validates the success of the fuzzy matching component.

- **Genre Filtering Effectiveness:**
  Filtering by genre significantly improved user satisfaction. When users looked for music in a particular mood (e.g., "lo-fi" or "pop") or books in a specific genre (e.g., "fantasy" or "romance"), the system effectively narrowed down relevant options, increasing precision.

- **Combined Feature Matching:**
  When users combined multiple filters (e.g., genre + rating + artist), the system maintained accuracy, showing that the hybrid logic combining content and metadata worked efficiently.

- **Cross-Domain Versatility:**
  Supporting both books and music in the same platform demonstrated the flexibility of the recommender logic. The architecture allowed for similar treatment of two distinct content types, implying the model could be extended to other media (e.g., movies or podcasts).

Interpretation of Implications:

- **User Empowerment:**
  The interactive filtering approach gives users more control over what they receive, which may lead to higher engagement and trust in the recommendations.

- **Content Discovery:**
  The system aids in content exploration. Users unfamiliar with certain genres or artists can still discover relevant content based on partial knowledge or preferences.

- **Foundation for Personalization:**
  While the current system is non-personalized, its solid performance with content features shows that integrating user history or behavior could result in even more powerful, personalized recommendations.

- **Feasibility for Real Deployment:**
  The results indicate that such a hybrid system is feasible for integration into larger platforms like book and music apps or learning tools, with real-time recommendations possible after optimization.

5. Discussion and Conclusion

5.1. Achievements

1. **Content-Based Filtering**: Using TF-IDF vectorization and cosine similarity, the system effectively identifies and recommends similar books or songs based on titles, authors/artists, genres, and other metadata.

2. **Filter-Based Search**: Users can filter recommendations by genre, author/artist, rating, or publication/release year, providing a flexible and user-friendly experience.

3. **Hybrid Recommendations**: By combining fuzzy title matching with TF-IDF and optional filters (e.g., genre and rating), the system increases recommendation accuracy and user satisfaction.

4. **Support for Two Domains**: The system is capable of recommending both books and songs, using separate but similar logic, allowing users to explore different types of content from a single interface.

5. **Interactive CLI Interface**: The command-line menu system provides an intuitive and clear way for users to interact with the recommender.

5.2. Limitations and Future Works

Cold Start Problem: The system does not support user profiles or historical behavior, making it less effective for personalized recommendations.

No Collaborative Filtering Yet: Recommendations are based purely on content; collaborative filtering like user-item matrix or user preferences could greatly improve personalization.

Scalability: With larger datasets, the TF-IDF and similarity calculations can become slow and memory-intensive. Optimization or use of approximate nearest neighbors could help.

No Web or Mobile Interface: The system currently runs in a terminal-based environment. A graphical user interface (GUI) or web application could enhance usability.

Data Quality and Enrichment: Some datasets may have inconsistent formatting, especially in genres or missing metadata. Cleaning and enriching the dataset would improve recommendation quality.

Reference & Source

Books_Dataset_GoodReads(May 2024). (n.d.). Kaggle. Retrieved April 29, 2025, from https://www.kaggle.com/datasets/dk123891/books-dataset-goodreadsmay-2024

KMeans — scikit-learn 1.6.1 documentation. (n.d.). Scikit-learn. Retrieved April 29, 2025, from https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

Music Dataset : 1950 to 2019. (n.d.). Kaggle. Retrieved April 29, 2025, from https://www.kaggle.com/datasets/saurabhshahane/music-dataset-1950-to-2019