Video Recommendation System

Abstract

This project presents a **Video Recommendation System** that leverages **content-based filtering** and **collaborative filtering** techniques to provide personalized video recommendations to users. By analyzing video metadata and user interactions, the system delivers tailored suggestions to enhance user engagement. The application is developed using **Streamlit** for the front end and integrates machine learning algorithms to ensure robust and accurate recommendations. This report details the development process, methodologies employed, system architecture, evaluation metrics, and future enhancements.

Project link

Demo Link

---

1. Introduction

1.1 Background

In the age of abundant digital content, users often face challenges in discovering videos aligned with their preferences. Recommendation systems bridge this gap by analyzing user behavior and content attributes, offering suggestions tailored to individual tastes.

1.2 Objectives

- To create a recommendation system that provides accurate and personalized video suggestions.
- To combine **content-based filtering** and **collaborative filtering** for enhanced recommendation quality.
- To provide an interactive and visually appealing user interface using Streamlit.

---

2. Features of the System

2.1 Recommendation Techniques

- **Content-Based Filtering**: Utilizes video metadata such as genre, description, and keywords to recommend videos similar to the user's previous choices.
- **Collaborative Filtering**: Employs user-item interaction data to recommend videos preferred by similar users.

2.2 User Interface

- Interactive video cards display:
  - Title
  - Genre
  - Description
  - Average rating with star visualization
  - Keywords as tags
- Mood-based filtering allows users to refine recommendations based on their mood (Relaxed, Neutral, Energetic).

2.3 System Evaluation

- Evaluation metrics include **Mean Absolute Error (MAE)**, **Root Mean Squared Error (RMSE)**, **Coverage**, and **Diversity**.
- Visual analytics feature a scatter plot comparing actual and predicted ratings.

---

3. System Architecture

3.1 Data Sources

- **Metadata**: Extracted from a CSV file (`summary.csv`) containing video details such as genre, description, and keywords.
- **User Interactions**: Derived from `user.csv` and `rate.csv`, representing user-item interaction data and ratings.

3.2 Data Pipeline

1. **Data Loading**: Import metadata and interaction data from CSV files.
2. **Preprocessing**:
   - Clean and normalize text data.
   - Vectorize text features using TF-IDF.
3. **Similarity Calculation**:
   - Compute cosine similarity for content-based filtering.
   - Construct a user-item interaction matrix for collaborative filtering.
4. **Recommendation Generation**:
   - Combine content-based and collaborative recommendations for final suggestions.

---

4. Methodology

4.1 Content-Based Filtering

- Uses **TF-IDF vectorization** to extract features from text fields (e.g., genre, description, keywords).
- Computes **cosine similarity** to identify similar videos.

4.2 Collaborative Filtering

- Constructs a **user-item interaction matrix** from rating data.
- Uses similarity measures to predict user preferences based on similar users.

4.3 Mood-Based Filtering

- Implements mood selection to refine recommendations by weighting genres and descriptions accordingly.

---

5. Implementation

5.1 Technologies Used

- **Programming Language**: Python
- **Front-End Framework**: Streamlit
- **Libraries**: `pandas`, `numpy`, `scikit-learn`, `scipy`, `plotly`

5.2 Installation

Clone the repository:
git clone <repository-url>

1. cd video-recommendation-system
2. Install dependencies:
   pip install -r requirements.txt
3. Prepare data files:
   - Place `summary.csv`, `user.csv`, and `rate.csv` in the directory.
4. Run the application:
   streamlit run app.py

---

6. System Evaluation

6.1 Metrics

- **Mean Absolute Error (MAE)**: Measures prediction accuracy.
- **Root Mean Squared Error (RMSE)**: Quantifies error magnitude.
- **Coverage**: Indicates the percentage of items for which recommendations are made.
- **Diversity**: Evaluates the uniqueness of recommended items.

6.2 Visualization

- Scatter plot comparing actual and predicted ratings highlights system performance.

---

7. Results

7.1 Accuracy

- The system achieved an MAE of 0.8 and an RMSE of 1.2, demonstrating reasonable prediction accuracy.

7.2 Coverage and Diversity

- Coverage: 85% of items were included in recommendations.
- Diversity: Recommendations included a wide range of genres and themes.

---

8. Future Enhancements

- **Deep Learning Integration**: Incorporate neural networks for better feature extraction and prediction.
- **Real-Time Updates**: Enable dynamic updates based on user interactions.
- **Multi-Language Support**: Expand the system’s accessibility.
- **Hybrid Recommendation Model**: Combine content-based, collaborative, and knowledge-based approaches.

---

9. Conclusion

The Video Recommendation System demonstrates a practical implementation of recommendation algorithms for enhancing user engagement on video platforms. By combining content-based and collaborative filtering with interactive features, the system offers a personalized and user-friendly experience. The integration of evaluation metrics and visual analytics provides insights into its performance and areas for improvement.