**Movie Recommendation System**

**Overview**

In this project, we developed a robust movie recommendation system using R and a rich dataset containing movie and user rating information. The system aims to provide personalized movie recommendations by analyzing factors that influence movie ratings and uncovering patterns in user preferences. This README file provides a comprehensive guide to the project's objectives, approach, analysis, and implications.

**Author**

Saron Yaya

**Initial work**

* [Portfolio Projects](https://github.com/Saron222/PortfolioProjects)

**Released on**

* GitHub

**My professional profile on LinkedIn**

* [My LinkedIn Profile](https://www.linkedin.com/in/saron-yaya/)

**Problem Statement**

The project addresses the following questions:

1. **What factors influence movie ratings?**
2. Are there underlying patterns in user preferences?
3. How can we build a recommendation system to offer personalized movie suggestions?
4. Do certain genres or movie release years significantly impact movie ratings?

**Approach**

**Data Processing**

1. **Cleaning and Transformation:**
   * Extracted and cleaned movie titles and release years.
   * Removed unnecessary characters and converted data types appropriately.
2. **Merging Datasets:**
   * Combined the movies and ratings datasets to create a unified dataset for analysis.

**Data Analysis**

1. **Summary Statistics:**
   * Provided descriptive statistics for the ratings and movies datasets to understand the data distribution.
2. **Visualization:**
   * Created histograms and boxplots to visualize rating distributions and movie release years.
3. **Correlation Analysis:**
   * Analyzed correlations between different variables to identify significant relationships.

**Recommendation System**

1. **Creating the Ratings Matrix:**
   * Transformed the user ratings data into a matrix suitable for collaborative filtering.
2. **Building the Recommender Model:**
   * Used User-Based Collaborative Filtering (UBCF) with cosine similarity to generate movie recommendations.

**Analysis**

**The analysis revealed several insights:**

* Distribution of Ratings: Highlighted viewer sentiments and genre popularity.
* Timestamp and Ratings Correlation: Indicated minimal impact of watch time on ratings.
* Nostalgia Effect: Older movies (pre-2000) generally received higher ratings.

**Implications for the Consumer**

The recommendation system enhances the movie-watching experience by:

* Providing personalized movie suggestions based on user preferences.
* Helping users discover new movies and avoid unsatisfactory choices.
* Aligning movie recommendations with user tastes.

**Limitations**

* Limited attributes and data size.
* Potential improvements with additional user demographics, contextual information, and advanced modeling techniques.
* Need for real-time interactions and feedback for a robust system.

**Conclusion**

A hybrid approach combining collaborative filtering and content-based filtering methods is promising. Future steps include refining techniques, implementing advanced algorithms, and incorporating sentiment analysis to improve recommendations further.

**Technologies Used**

* R
* tidyverse for data manipulation
* ggplot2 for data visualization
* recommenderlab for building recommendation models

**Installation**

1. Clone the repository: git clone [https:// https://github.com/Saron222/PortfolioProjects.git](https://github.com/saronyaya/DSC510.git)
2. Navigate to the project directory: **cd Movie Recommendation System.R**

**Contributing**

1. Fork the repository from [Saron222/PortfolioProjects](https://github.com/Saron222/PortfolioProjects/fork)
2. Create your feature branch: **git checkout -b feature/your-feature-name**
3. Commit your changes: **git commit -am 'Add your feature'**
4. Push to the branch: **git push origin feature/your-feature-name**
5. Create a new Pull Request in the original repository