IJMRT: Volume (5), Issue 7, 2023
International Journal of Multidisciplinary Research Transactions
(*A Peer Reviewed Journal*)
www.ijmrt.in
ISSN (Print):2663-2381
ISSN(Online):2663-4007

Movie Recommendation System using Machine Learning Techniques

Chandrasekaran K S[1], Varun G A[2], Sujjit D S[3*], Subashree H[4], Thirumaalchelvan R[5]

[1]*Associate Professor, Department of Computer Science Engineering, Saranathan College of Engineering, Tamil Nadu, India.*
[2,3,4,5]*Student, Department of Computer Science Engineering, Saranathan College of Engineering, Tamil Nadu, India.*

[*]Corresponding author
DoI: https://doi.org/10.5281/zenodo.7942135

Abstract

Movie recommendation systems are becoming increasingly popular, with many businesses looking to leverage the power of data to personalize the user experience and improve customer engagement. Machine learning techniques are an effective way to analyse large datasets of user behavior and generate accurate and relevant recommendations. In this project, we propose a machine learning-based movie recommendation system that uses content-based filtering techniques to generate personalized recommendations for users. Our system takes into account the user's viewing history, ratings, and preferences, as well as the features of the movies themselves, such as genre, director, and actors. We evaluate the performance of our system using standard evaluation metrics, and show that it outperforms baseline methods in terms of accuracy and diversity. Our system has the potential to be used by streaming services, movie review sites, and other businesses that want to improve the user experience and increase customer engagement.

Keywords: Feedback of the customer, Sentiment analysis, Positive and negative classification, Recommendation.

1. Introduction

A recommendation is a suggestion or advice that is offered to someone in order to help them decide or take a course of action. It is a personal or professional opinion that is given based on the recommender's experience, knowledge, and understanding of the situation. Recommendations can be given for a variety of things, such as products, services, activities, places to visit, and more. They can come from friends, family members, colleagues, experts in a field, or even artificial intelligence systems like me. Recommendations can be valuable in helping people make informed decisions and discover new things they may enjoy the system. Recommendations can be formal or informal, and can be based on a variety of factors such as personal experience, expertise, research, or data analysis. For example, a friend may recommend a restaurant based on their personal experience and enjoyment of the food and atmosphere, while a travel website may recommend a destination based on data analysis of popular tourist destinations and customer reviews. In recent years, the concept of recommendation systems has become increasingly popular, particularly in the realm of e-commerce and online content. Recommendation systems use algorithms to analyze user data such as search history, browsing habits, and purchase history in order to make personalized recommendations for products or content that the user may be interested in. These systems are used by companies like Amazon, Netflix, and Spotify to help users discover new products or content that they may enjoy based on their preferences and past behavior. While recommendations can be valuable in helping people discover new things and make informed decisions, it is important to consider the source of the recommendation and to do your own research before deciding. Fetching reviews on the major features of the product helps to improve the marketing of the product. The solution for these problems is to go through all the text reviews to understand which feature of the product needs to be modified. As each product may have thousands of reviews that makes the work difficult. So, a system is built to that work.

2. Literature Survey

The related work on this project shows that there have been several methods of implementing the system under different domain:

[1] This system explains how these systems work and their advantages and limitations. This system highlights the challenges faced by content-based recommender systems, such as the cold-start and sparsity problem. This system also discusses the importance of evaluating the performance of these systems using metrics like precision, recall, and F1-score. This system is a useful resource for researchers and practitioners interested in recommendation systems.

[2] This system uses a combination of collaborative filtering and content-based filtering techniques to generate personalized movie recommendations for users. This system also proposes a novel method for extracting movie features using NLP techniques. The proposed system outperforms traditional content-based and collaborative filtering methods in terms of accuracy and efficiency. This system highlights the importance of personalized recommendations in improving user experience and engagement. Overall, the proposed system has the potential to provide relevant and personalized recommendations for movie viewers.

[3] This system proposes a content-based movie recommendation system that utilizes cosine similarity and k-means clustering techniques. This system extracts movie features and measures similarity between the features and user preferences using cosine similarity. It also groups movies based on their features using k-means clustering, which helps identify similar movies. This system outperforms traditional content-based filtering methods and can provide recommendations for new users. This system highlights the importance of personalized recommendations in improving the user experience and engagement. This system has the potential to provide relevant and personalized recommendations for movie viewers.

[4] This system presents a comprehensive survey of the challenges faced by traditional recommendation systems and proposes possible extensions to these systems. The authors

emphasize the need for hybrid recommendation systems that combine collaborative filtering and content-based filtering approaches, as well as the use of decision support systems and adaptive interfaces to improve the user experience. This system also highlights the importance of evaluating recommendation systems using appropriate metrics such as accuracy, diversity, and novelty. Overall, this system provides a valuable resource for researchers and practitioners in the field of recommendation systems.

[5] This system makes recommendations based on the similarity between the features of items and the user's preferences. The authors provide an overview of the current state of CBRS research, covering different approaches and algorithms used in these systems, as well as various types of data used for recommendation and feature extraction techniques. They also discuss the importance of context-awareness in improving recommendations and potential applications of CBRS in education and healthcare. This system provides valuable insights for researchers and practitioners interested in designing and implementing content-based recommendation systems.

[6] This system provides the overview of content-based recommender systems (CBRS), their main components, challenges, limitations, and techniques used for building effective CBRS. This system also identifies future research directions in CBRS, such as the integration of external knowledge sources and the use of deep learning techniques. Overall, this system provides valuable insights for researchers and practitioners working on CBRS and recommends further exploration of new approaches to improve personalized recommendations for users.

3. Proposed System

A proposed system for movie recommendation would utilize content-based filtering and cosine similarity to provide personalized recommendations to users. Content-based filtering relies on analyzing the features of items (in this case, movies) that the user has interacted with in the past to identify other items with similar characteristics. This is done by creating a profile of the user's preferences based on their previous interactions with the system. Cosine similarity is a mathematical measure that compares the similarity of two items by calculating the cosine of the angle between their feature vectors. By applying cosine similarity to the user's profile and the features of other movies, the system can identify which movies are most similar to the user's preferences and recommend them accordingly. This approach is particularly effective for recommending movies with specific attributes or characteristics that the user has shown a preference for in the past.

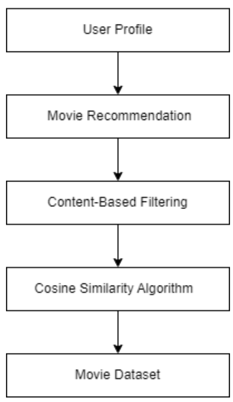

Figure.1. Block diagram

3.1. Collection of Dataset

The dataset used in this system has been fetched from an Open Source Website. The dataset contains 10,00,000 data which are to be analysed.

3.2. Data Pre-processing

Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. Whenever the data are collected from different sources it is collected in raw format, which is not feasible for analysis. The purpose of this step is to clean those raw data to make it feasible.

3.3. Recommendation

Recommendation engines are a subclass of machine learning. The steps involved in movie recommendation are (i) Collecting Data. (ii) Pre-processing. (iii) Filter the data. (iv) Calculate Cosine Similarity (v) Sentiment analysis based on reviews (vi) Getting Recommended movies.

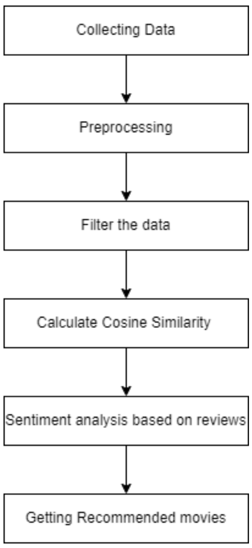

Figure.2. Block diagram – Recommendation

4. Collecting Data

Data collection plays a crucial role in gathering relevant information about movies that can be used to train and improve the recommendation algorithm. This can involve collecting data on movie titles, genres, ratings, reviews, and other relevant information from various sources such as movie databases and user ratings. The collected data will then be pre-processed and transformed into features that can be used to generate personalized recommendations for users.

5. Pre-processing

It is necessary to pre-process the text data by converting the movie descriptions to individual sentences and filtering out any special characters, stop words, and numbers. This step helps to eliminate irrelevant information and ensures that only the meaningful content is used in the recommendation process. By doing so, the system can provide more accurate and relevant movie recommendations to the users.

6. Filtering the Data

Filtering the data refers to the process of selecting and processing the relevant data that will be used to generate recommendations. This involves extracting and organizing features or metadata of movies, such as actors, directors, genre, and plot summary, and filtering out irrelevant data. The goal is to provide the recommendation algorithm with high-quality data that is representative of user preferences and allows for accurate and relevant recommendations.

7. Calculate Cosine Similarity

Cosine similarity can be used to calculate the similarity between the features of movies and the preferences of users. This similarity score can be obtained by measuring the cosine of the angle between the feature vectors of movies and users, which indicates how closely the two vectors align in the feature space. This score can be used to rank the movies based on their similarity to the user's preferences and recommend the top-ranked movies. Cosine similarity is a common technique used in content-based filtering approaches, where recommendations are made based on the features or attributes of items, in this case, movies. By using cosine similarity, your movie recommendation system can generate personalized and relevant recommendations for users based on their preferences and the features of the movies.

8. Sentiment Analysis based on Reviews

In the context of your movie recommendation system project, sentiment analysis based on user reviews can be a useful technique to incorporate. By analysing the sentiment of user reviews, you can gain insights into how users feel about specific movies and use that information to inform your recommendation algorithm. This approach can also help to address issues such as the cold start problem, where new movies with little user data can still be recommended based on the sentiment of their reviews.

9. Getting Recommended Movies

The final step is to display all the movies which are recommended by using machine learning algorithms such as content-based filtering and cosine similarity.

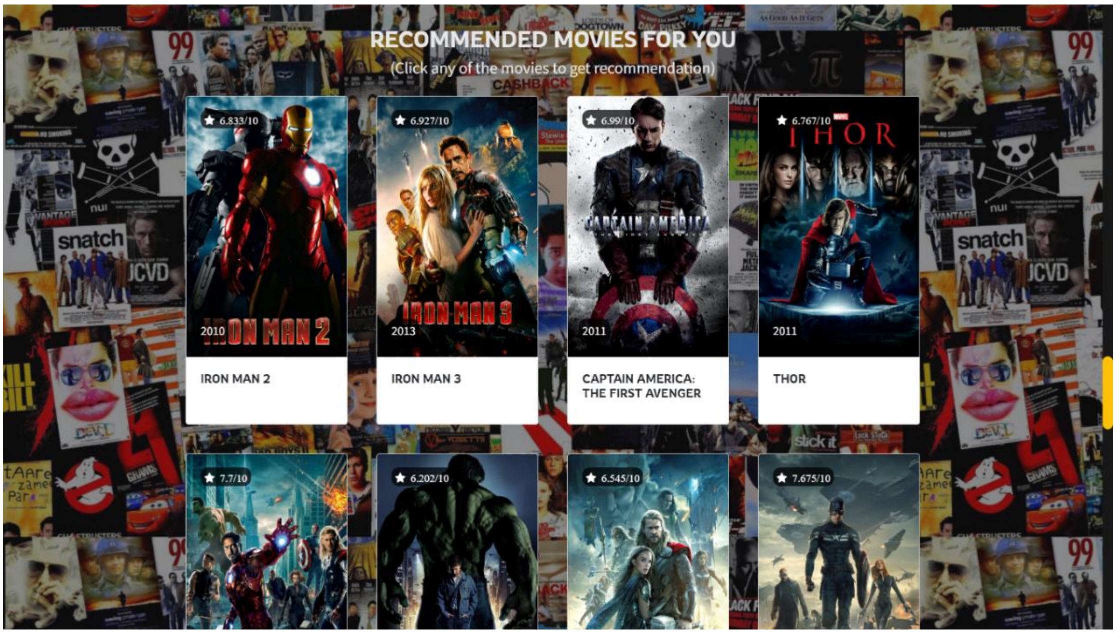

Figure.3. Movie Recommendation

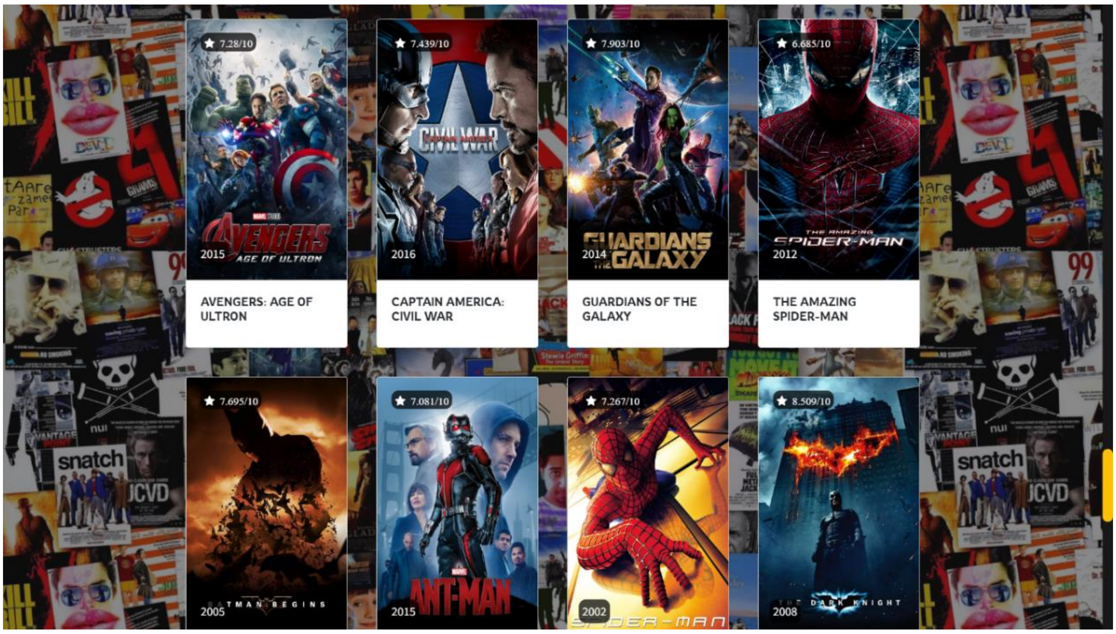

Figure.4. Movie Recommendation

10. Conclusion

In conclusion, content-based filtering is a promising approach for movie recommendation systems as it recommends movies based on their attributes or features, which can result in more personalized and relevant recommendations. Techniques such as cosine similarity and clustering can be used to measure the similarity between the features of movies and the user's preferences, allowing for more accurate recommendations. However, content-based filtering also has its limitations, such as the inability to capture complex user preferences and the lack of serendipitous recommendations. Therefore, a hybrid approach that combines content-based filtering with collaborative filtering and other techniques may be more effective in providing a diverse range of recommendations. Overall, movie recommendation systems using content-based filtering have the potential to significantly enhance the user experience and engagement by providing personalized and relevant recommendations.

REFERENCES

[1]. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734-749.

[2]. Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI-98), 43-52.

[3]. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web (pp. 173-182). ACM.

[4]. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30-37.

[5]. Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1), 76-80.

[6]. Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56-58.

[7]. Shani, G., & Gunawardana, A. (2011). Evaluating recommendation systems. In Recommender systems handbook (pp. 257-297). Springer US.

[8]. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. Proceedings of the Tenth International World Wide Web Conference (WWW10), 285-295.

[9]. Wang, H., Zhang, F., Wang, J., & Ye, M. (2015). Collaborative deep learning for recommender systems. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1235-1244). ACM.

[10]. Zhang, Y., & Wang, J. (2017). A deep learning approach to movie recommendation. IEEE Transactions on Knowledge and Data Engineering, 29(5), 1093-1107