ISSN: 2249-7196
IJMRR/April-June. 2025/ Volume 15/Issue 2s/510-518
Syed Ibrahim/ International Journal of Management Research & Review

MOVIE RECOMMENDER SYSTEM USING SENTIMENT ANALYSIS.

Syed Ibrahim[1], Mohammed Irfan Ahmed[2], Mohammed Faizur Rahman[3], Dr. K.Nagi Reddy[4]
[1,2,3]B.E. Student, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
[4]Head of Department, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
k.nagireddy@lords.ac.in

ABSTRACT

In today's digital landscape, users are constantly bombarded with countless options on platforms such as streaming services, e-commerce sites, and social media. This sheer volume of content can lead to decision fatigue, making it difficult for users to discover content that truly resonates with their interests. To alleviate this challenge, Recommendation Systems (RS) have become pivotal in delivering personalized suggestions, thereby improving user satisfaction and engagement. Conventional recommendation strategies typically fall into two primary categories:

Content-Based Filtering (CBF): Suggests items similar to those a user has previously interacted with, using attributes like genre, director, or cast. While effective in personalization, CBF can suffer from limited diversity—often referred to as the "over-specialization" issue. Collaborative Filtering (CF): Relies on analysing user interactions, such as ratings and reviews, to identify patterns and recommend items enjoyed by similar users. However, CF faces challenges like the cold-start problem, where recommendations are hindered due to insufficient data on new users or items.

Keywords: Recommendation Systems, Sentiment Analysis, Ensemble Learning, KNN Algorithm, Collaborative Filtering, Content-Based Filtering, Hybrid Systems, User Experience.

I. INTRODUCTION

The explosion of digital content in recent years, users often face a dilemma when choosing what to watch, read, or buy. This information overload makes it harder to find content that aligns with individual preferences. Recommendation Systems (RS) have emerged as a crucial solution, streamlining the discovery process by suggesting content tailored to user tastes.

These systems are now a staple across a range of domains from online shopping (Amazon), music platforms (Spotify), and

travel services to movie streaming platforms like Netflix and Prime Video.

In the realm of movie recommendations, two common techniques are widely used [1]

Content-Based Filtering (CBF): Recommends content based on the features of movies (such as genre, actors, and director) that the user has shown interest in before [2].

Collaborative Filtering (CF): Suggests items by analysing the preferences and behaviours of users with similar tastes. While each method is effective on its own, both come with limitations [3].

CBF tends to recommend similar content repeatedly, resulting in a lack of variety. On the other hand, CF is heavily dependent on user data and struggles when there is little or no data available a problem known as the "cold start" issue [4].To overcome these shortcomings, hybrid recommendation systems have gained popularity. These systems combine multiple algorithms to deliver more balanced and accurate recommendations.

The proposed system enhances the hybrid approach further by incorporating sentiment analysis. By evaluating user reviews and emoji responses, the system can better understand the user's emotional response to content. This emotional layer of understanding leads to more engaging and relevant suggestions.

To address these shortcomings, this paper introduces a hybrid movie recommendation system that combines the strengths of multiple techniques:

- K-Nearest Neighbour's (KNN)-Based Collaborative Filtering: Utilizes cosine similarity to determine relationships between users or items, allowing for behaviour-driven recommendations based on historical data [5].
- Content-Based Filtering: Leverages metadata including genre, cast, and director to recommend content aligned with a user's previous interests.
- Sentiment Analysis: Applies Natural Language Processing (NLP) to interpret user-generated reviews and emoji feedback, assigning sentiment scores that capture emotional responses. This emotional intelligence enriches the relevance of recommendations [6].

The model employs an ensemble learning approach to merge the outputs from CF, CBF, and sentiment analysis. By fine-tuning the weight of each component, the system ensures a well-rounded and tailored recommendation experience.

The system is validated using a real-world dataset enhanced with user ratings, textual reviews, and emoji reactions. The results

indicate that our hybrid model significantly outperforms traditional methods across metrics such as Precision, Recall, F1-Score, and Root Mean Square Error (RMSE) [7]. Notably, integrating sentiment analysis substantially improved user satisfaction by incorporating emotional depth into the recommendation logic.

II. RELATED WORK

Recommendation systems have evolved considerably, with various improvements in machine learning and natural language processing enabling smarter and more personalized recommendations. There are so many techniques that has been already studied about the recommender system. Some of them is based on amount of weighting of the data and other as based on user interest. There are so many algorithms already developed for this that will reduce the time of the user and difficulty level.

Sentiment Analysis in Recommendation Systems

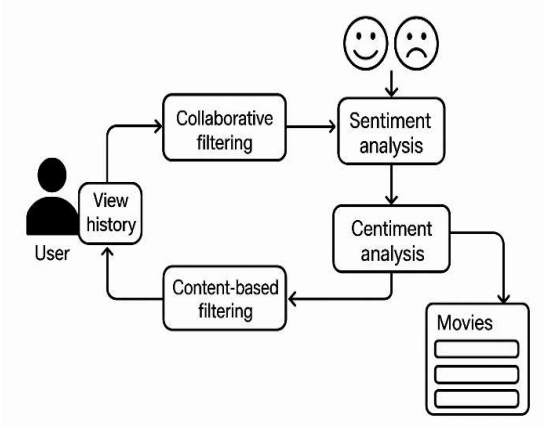


Fig.1. Work-flow with CF, CBF, and SA

It requires a lot of prior knowledge on the basic of rating of the movie given by user. It mainly uses movie datasets for evaluation and testing purpose. But the algorithm, system designed are not working efficiently, but the study is in process to resolve this issue and make the system more perfect and accurate. Collaborative Filtering and Content based Filtering were implemented.

Collaborative Filtering (CF):Collaborative Filtering (CF) stands as one of the most foundational and extensively used techniques in the field of recommendation systems. It is renowned for its simplicity, effectiveness, and ability to produce highly personalized suggestions without needing deep domain knowledge about the items being recommended. The core idea behind collaborative filtering is to identify patterns in user behaviour, particularly focusing on how users rate or interact with various items—such as movies, books, or products—and using this data to uncover similarities between users or items.

Unlike content-based approaches that rely on item metadata (such as genre, cast, or director in the case of movies), collaborative filtering solely depends on user-item interaction data [8].

This makes CF particularly powerful in scenarios where rich metadata is unavailable or inconsistent across the

dataset. CF techniques assume that if users agreed in the past, they will likely agree in the future too. Hence, recommendations are generated by identifying these patterns of agreement.

There are two main categories within collaborative filtering:

- User-Based Collaborative Filtering
- Item-Based Collaborative Filtering

1. User-Based Collaborative Filtering

In user-based CF, the system identifies users who have similar tastes or preferences by comparing their historical rating patterns. The intuition here is straightforward: if User A and User B both rated a set of movies similarly—say, they both enjoyed action and thriller genres—it's likely they'll agree on future preferences too. So, if User A has rated a movie that User B hasn't seen yet, the system might recommend that movie to User B, assuming shared interests.

The steps typically involve:

Constructing a user-item matrix where rows represent users and columns represent movies.

Computing similarity scores between users using metrics such as cosine similarity or Pearson correlation.

Selecting the top 'K' most similar users (K-nearest neighbours).

Generating a predicted score for unrated movies based on the preferences of these neighbours.

Item-Based Collaborative Filtering:

Item-based CF flips the perspective. Instead of finding similar users, it focuses on identifying similarities between items themselves. For example, if a user has watched and liked "Inception", and many users who liked "Inception" also liked "Interstellar", then "Interstellar" might be recommended to the user[9].

This method is especially useful in systems with a large number of users because item-item similarities tend to remain more stable over time compared to user-user relationships, which can be more dynamic.

In this approach:

The system computes item similarity using co-occurrence or correlation in user ratings.

Recommendations are generated by finding items that are most similar to those already rated highly by the user.

This strategy is often faster and scales better for large datasets.

Content-Based Filtering (CBF):

Content-based filtering focuses on item features. In movie recommendations, attributes such as genre, cast, and director are used to match content with user preferences. It is particularly useful for new users since it doesn't rely on the behaviour

of other users. However, it can become repetitive and limit content diversity.

Content-based Filtering works with a motive to use the feature of the dataset and recommend the suggestion same as the items that are close and parallel to the user's likes and dislikes based on their past action and their rating on the particular items.

Sentiment Analysis in Recommendation Systems:

Traditional recommendation systems primarily rely on user behaviour such as ratings, clicks, or watch history to suggest relevant content. While these systems have proven effective, they often overlook a critical dimension of user feedback: emotion. This is where sentiment analysis enters the picture, introducing a layer of emotional intelligence into the recommendation process for which the expression used is FinalScore=a.CF Score + (1-a).SentimentScore

Sentiment analysis, also known as opinion mining, is a Natural Language Processing (NLP) technique that involves analysing and interpreting the emotional tone behind textual or symbolic data [10]. In the context of recommendation systems, sentiment analysis helps in understanding how users feel about the content they have interacted with not just what they rated or watched. This emotional understanding allows the system to deliver recommendations that are not only behaviourally relevant but also emotionally aligned with user preferences[11].

Emotional Intelligence in Recommendations:

Emotional intelligence (EI), also known as emotional quotient (EQ), is the ability to understand, manage, and use emotions effectively.

It involves recognizing and understanding your own emotions and the emotions of others[12], and using that information to guide your thinking and behaviour.

When a user leaves a review or comment like "This movie was absolutely breath-taking!" or simply uses emoji's such as "😍" or "😭", they are offering valuable insights into their emotional engagement with the content. Sentiment analysis captures this qualitative data and translates it into structured sentiment scores that the system can utilize.

This means the recommendation engine doesn't just record that the user watched a particular movie it also learns that the user loved, liked, or even disliked it.

This emotional context makes a significant difference. For instance, two users may have watched the same five movies, but if one rated them highly while the other expressed dissatisfaction in their comments, a system enhanced with

sentiment analysis will distinguish their preferences more accurately.

Without sentiment analysis, both users might receive the same future recommendations, leading to a subpar user experience.

III. METHODOLOGY

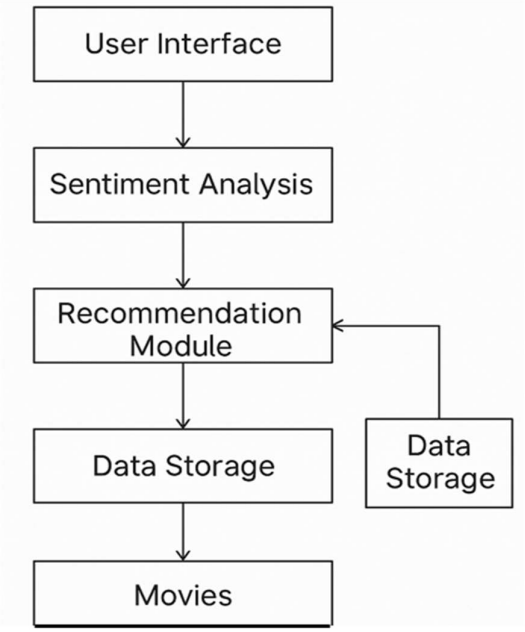
Our hybrid model combines three core components: collaborative filtering, content-based filtering, and sentiment analysis.

Collaborative Filtering using KNN

We apply the K-Nearest Neighbour's (KNN) algorithm to identify similar users or items.

Steps involved:

1. User-Item Matrix Construction: Users and movies form a matrix where each cell represents a rating.



2. Cosine Similarity: Measures how closely aligned user preferences are using the process flow.

Fig.2. Process Flow for Sentiment Aanalysis

3. Neighbour Selection: Picks the top-K similar users or items.
4. Rating Prediction: Estimates ratings using neighbours' input.

$$\text{sim}(u,v) = \frac{\sum_{i=1}^{n} r_{u,i} \cdot r_{v,i}}{\sqrt{\sum_{i=1}^{n} r_{u,i}^2} \cdot \sqrt{\sum_{i=1}^{n} r_{v,i}^2}}$$

Fig.3. Equation implemented to find sim(u,v)

Integrating Sentiment Scores

To add emotional intelligence, we analyse user reviews and emoji's:

Text Review Classification: NLP techniques tag reviews as positive, neutral, or negative.

Emoji Mapping: Emoji's are converted into sentiment scores.

Final Score is calculated using a weighted average:

Where α alpha controls the influence of collaborative filtering vs. sentiment.

System Architecture:

The architecture is modular and designed for scalability:

- Frontend Interface: Allows users to browse, rate, and review movies.
- Database Layer: Stores metadata, user interactions, and sentiment scores.

- Data Pre-processing Unit: Cleans and tokenizes text, maps emoji's.
- Sentiment Analyser: Uses NLP models to detect sentiment polarity.
- Recommendation Engine: Combines CBF, CF, and sentiment to generate suggestions.
- Backend Server: Manages APIs, authentication, and communication layers.

IV.RESULTS AND DISCUSSSION

Dataset: We used the Movie Lens dataset, enriched with scraped reviews and emoji reactions from online platforms.

Dataset Details:

100,000+ ratings from over 5,000 users

10,000+ reviews with sentiment tags

Movie metadata including genre, cast, and director

Evaluation Metrics To Assess Performance:

Precision: Proportion of relevant items among recommended ones

Recall: Proportion of relevant items successfully recommended

F1-Score: Balance between precision and recall

RMSE: Measures deviation between predicted and actual ratings

Evaluated Model Comparison

1. Baseline CF model
2. CBF-only model
3. Hybrid model with Sentiment Analysis

Performance Results:

Precision: 0.84 Recall: 0.79

F1-Score: 0.81 RMSE: 0.92

The hybrid model demonstrated superior performance and deeper personalization.

V.CONCLUSION

In today's digital landscape, where users are overwhelmed by an abundance of content across streaming platforms, finding the right movie to watch can be both time-consuming and frustrating. This study addresses that challenge by proposing a hybrid movie recommender system that not only learns from user behaviour but also interprets their emotions, offering deeply personalized suggestions.

The proposed model integrates three key techniques Collaborative Filtering (CF), Content-Based Filtering (CBF), and Sentiment Analysis into a unified recommendation engine. The collaborative filtering component, powered by the K-Nearest Neighbour's (KNN) algorithm, identifies patterns in user interactions and similarities among users and items. The content-based module draws upon metadata such as movie genre, cast, and director to align recommendations with the user's established preferences. However, what sets this model apart is the addition of

sentiment analysis, which interprets the emotional tone behind user reviews and emoji reactions. By incorporating Natural Language Processing (NLP) techniques, the system can analyse textual reviews and map emoji's to sentiment scores, allowing it to capture the user's emotional responses to previously watched content. These sentiment scores are then integrated with the collaborative filtering predictions through a weighted scoring mechanism, which allows the model to balance between objective interaction data and subjective emotional feedback.

In addition to its effectiveness, the architecture of the system is modular, scalable, and domain-agnostic, meaning it can be adapted beyond movie recommendations to other sectors like music, e-commerce, education, and online shopping. This flexibility allows developers and researchers to extend the concept of emotionally intelligent recommendations into a wide array of digital services.

This sentiment-aware hybrid approach not only improves the accuracy of recommendations but also enhances user satisfaction by tailoring results to both behaviour and emotional context. The system was validated using a comprehensive real-world dataset enriched with ratings, reviews, and emoji feedback. Performance metrics such as Precision, Recall, F1-Score, and Root Mean Square Error (RMSE) indicate that the hybrid model outperforms conventional methods, particularly in capturing the nuanced preferences of users.

VI. REFERENCES

[1] Golbeck, J. (2006). Generating Predictive Movie Recommendations from Trust in Social Networks. LNCS, vol. 3986, pp. 93–104.

[2] Mohapatra, H., Panda, S., Rath, A., Edalatpanah, S., & Kumar, R. (2020). A tutorial on PowerShell pipeline and its loopholes. International Journal of Emerging Trends in Engineering Research, 8(4), 975-982.

[3] Xu, X. (2012). From cloud computing to cloud manufacturing. Robotics and Computer-Integrated Manufacturing, 28(1), 75-86.

[4] Mohapatra, H. (2009). HCR using neural network (Doctoral dissertation, BPUT).

[5] Yan, B., & Chen, G. (2011). AppJoy: Personalized mobile application discovery. MobiSys 11, ACM.

[6] Manikrao, U. S., & Prabhakar, T. V. (2005). Dynamic selection of web services with recommendation system. NWESP'05. IEEE.

[7] Von Reischach, F., et al. (2009). A mobile product recommendation system. IEEE Pervasive Computing, pp. 1-6.

[8] Choi, K., et al. (2012). A hybrid online-product recommendation system. Electronic Commerce Research and Applications, 11(4), 309-317.

[9] Goldberg, D., et al. (1992). Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12), 61-70.

[10] Bell, R. M., & Koren, Y. (2007). Scalable collaborative filtering. ICDM 2007. IEEE.

[11] Arora, G., et al. (2014). Movie recommendation system based on users' similarity. IJCSMC, 3(4), 765-770.

[12]Bhatt, B., et al. (2014). A review paper on ML-based RS. IJEDR, 2(4), 3955-3961.