Journal of Global Integrated Research

Vol. 01 No. 01
DOI: XYZ
ISSN: XYZ

Enhancing Movie Recommendation Systems Using Hybrid Filtering and Twitter-Based Sentiment Analysis

K.Shanmuka Sai[1], Ch.Anil[1], B.Naga Vardhan[1], G.Sai Bhargav[1] , A.Yaswanth Kumar[2*]

[1]Department of Information Technology, R.V.R&J.C College of Engineering, Chowdavaram,Guntur, Andhra Pradesh-522019. Email Id: shanmukasaikattamudi@gmail.com

[1]Department of Information Technology, R.V.R&J.C College of Engineering, Chowdavaram,Guntur, Andhra Pradesh-522019. Email Id: anilchintha658455@gmail.com

[1]Department of Information Technology, R.V.R&J.C College of Engineering, Chowdavaram,Guntur, Andhra Pradesh-522019. Email Id: bandaruvardhan90@gmail.com

[1]Department of Information Technology, R.V.R&J.C College of Engineering, Chowdavaram,Guntur, Andhra Pradesh-522019. Email Id: saibhargavgudapati8@gmail.com

[2*]Department of Information Technology, R.V.R&J.C College of Engineering, Chowdavaram,Guntur, Andhra Pradesh-522019. Email Id: aykumar@rvrjc.ac.in

Corresponding Author's :-A.Yaswanth Kumar, Email: aykumar@rvrjc.ac.in

Received:-28/01/2026, Revised:-10/03/2026, Accepted:-17/03/2026, Published:-24/03/2026

ABSTRACT

Recommendation Systems (RS) have become an integral component of the online world, assisting users in navigating through the extensive information space of online e-commerce, as well as online entertainment systems.Although methods such as Collaborative Filtering (CF) and ContentBased Filtering (CBF) have been shown to be efficient and effective, they suffer from the problem of requiring a large historical user preference profile, thus often facing challenges such as the Cold Start Problem and Sparsity. To address such limitations, a multimodal movie recommendation framework that combines CF, CBF, and Sentiment Analysis of microblogging characteristics, namely, tweets, is presented in this paper. The addition of the Sentiment component ensures that the system encodes real-time public opinion, trending, and immediate reactions of the audience, thus enhancing the recommendation space beyond just rating matrices.This proposed framework uses the MovieTweetings dataset, which contains user ratings and various pieces of movie information, making it an ideal choice for sentiment-based analysis. This is because all tweets are processed and purified from noise consisting of slang terms, hyperlinks, and special characters before being scored using the VADER analyzer tool. Compound scores are then converted into a numerical rating system and combined using weighted scores of similarity indices consisting of various pieces of movie information. Experimental analysis proves the correlation between the ratings and IMDb ratings, and this proves the efficiency and effectiveness of combining the social signals. Correlation factors like PLCC, SROCC, and KRCC show the alignment of the sentiments extracted from the tweets and the preferences of the users. In addition to this, the hybrid model performs better on the Top-N recommendation procedure than the standard CF and CBF approaches. This approach generally establishes the significance of combining the structured ratings and the unstructured data from the social media site.

Keywords: *Movie Recommendation System; Collaborative Filtering; Content-Based Filtering; Sentiment Analysis; Twitter; VADER*

1. Introduction

Today, with the advent of the digital age, the Internet has become an essential component in the day-to-day life of a person, providing instantaneous access to a vast amount of information in different fields for the Internet user. Though quick growing need for information on the Internet has increased its accessibility, the very same phenomenon has given rise to the problem of information overload. The Internet user is faced with the uphill task of extracting relevant information out of the enormous quantum of available data. Herein lies the need to extract relevant information with the help of Recommendation Systems (RS).

Recommendation systems have seen widespread adoption in a wide range of application fields, including e-commerce websites, online learning platforms, social networking sites, and online entertainment. Among these, movie recommendation systems have gained much attention due to the extensive employment of online entertainment platforms such as Netflix, Amazon Prime Video, Disney+, and IMDb. Movies are a primary source of joy and fun for users, who have begun making heavy use of intelligent recommendation platforms to discover new sources of their interest.

Movies can be described based on various parameters, such as genre, language, year of release, director, producer, and cast. Current online movie platforms try to provide personalized experiences to users based on their historical data regarding watch history, ratings, and browsing behavior. Based on such data, a movie recommendation system is designed to offer recommendations of movies that are very similar to what a particular user likes. But these systems are largely dependent on how accurate this data is.

The traditional methods of movie recommendation are generally grouped into two classes, namely: Collaborative Filtering (CF) and Content-Based Filtering (CBF). In the Collaborative Filtering method, similarities among users and/or objects are tried to be obtained. It relies on the basic principle that those people who have similar preferences in the past will also have similar preferences in the future. In the case of the Content-Based Filtering method, a movie is recommended depending on comparisons done among attributes, which are then matched with a preference profile owned by a user. Such methods have proven their worth in controlled lab settings.

Although CF and CBF methods have gained popularity, these algorithms face numerous challenges. In collaborative filtering, there exist sparse data problems due to a high number of users and items with fewer interactions between them. Another problem that may arise is the cold-start problem whereby there may exist new users or new movies with fewer rating data, which may generate incorrect recommendations. Another problem with Content-Based Filtering is that it may cause overspecialization where users can be presented with items having similar properties.

However, in order to handle these issues, some research studies have focused on incorporating sources other than the rating matrix. Social networking sites are recently gaining momentum as a great source of user-generated information. Social networking sites, such as Twitter, Facebook, and Instagram, provide users with the facility for free expression of opinion, feelings, and experiences regarding movies and other entertainment content on the site. Among these, Twitter has become very popular due to real-time functionality and public availability.

Twitter enables subscribers to send small messages viewed as tweets, which may include quick reactions, public opinions, and trending views on films. Unlike film rating systems, tweets enable users to give quick responses that instantly reach the perception of viewers who come out of the theatres following the release of films, with Twitter emerging as the perfect stage regarding gaining insight into public trends and awareness of people about films. Sentiment analysis is basically a computer technique that tries to extract emotional polarity from unstructured data sources, such as texts, to identify if their nature is positive or negative with a certain degree of accuracy. By carrying out sentiment analysis on movie-related tweets, it is possible to deduce audience opinion and hence incorporate that opinion into a movie recommendation system. Encouraged by such findings, this presented research work aims at enhancing movie recommendation systems with the incorporation of hybrid filtering approaches and Twitter sentiment analysis. The approach integrates Collaborative Filtering and Content-Based Filtering methods, each drawing their benefits from user behavior and movie-related features, respectively. Furthermore, the sentiment scores retrieved through the VADER analyzer will be utilized in establishing the current opinions and reactions of audiences on the displayed films. By integrating the two sources, the system enhances the Top-N accuracy film recommendation problem. The efficiency of the proposed method ascertains through correlation analysis and recommendation performance measures. The analysis reveals that the incorporation of social media sentiment information into traditional recommendation methods results in more accurate movie recommendations. Of importance, this research work unveils the importance of social signals toward modern recommendation methods. It unveils how the inclusion of social signals may reduce the effects associated with a sparse data environment.

2. Related Work

Recommendation systems have been widely researched over the years due to the need to manage the problem of information overload that pervades all online platforms. The most widely used techniques among the ones that were proposed first, and which provide the cornerstone of all modern-day systems, are Collaborative Filtering (CF) and Content-Based Filtering (CBF).CF operates by finding relationships between users, which may be within items, so that versions of the past can be utilized to make appropriate recommendations. Though this has

proved successful, there still exist several difficulties that namely relate to the sparseness, scalability, and cold start problems, especially when new users and new items are added into a system.

Content-Based Filtering concentrates on the process of recommending items based on their intrinsic properties that align with the user preference profile. In the movie recommendation domain, Content-Based Filtering relies on movie metadata that includes genre, cast, director, language, and release year. This filtering technique may result in decreased reliance on other users' information, which may diminish privacy issues associated with other filtering techniques. However, it may cause overspecialization because users may continuously be exposed to items that align with items they have already used before.

To address the limitations of CF and CBF individually, there has been research on creating hybrid systems for CF. A hybrid system is a combination of two or more methods to take advantage of their individual strengths and compensate for their inadequacies. There are various methods for hybridization. These methods include combination, switching, and fusion. Many researchers have concluded that hybrid systems are more effective than single method-based systems, especially when handling a sparse data population. However, current hybrid CF systems are also dependent on ratings and are less capable for newly added materials.

Recently, there has been interest in the community in utilizing auxiliary sources of information for improving the accuracy of recommendations. Context and social-aware recommendation systems gather information about the context in which the recommendation is being made. These include time, location, and social relationships. Socialaware systems make use of the assumption that people related through social contacts have common interests. Although social-aware systems improve personalization techniques, these systems are dependent on explicit social connections for their accuracy.

The rising popularity of social media sites has generated potential opportunities for recommendation systems, given the availability of user-generated content on a large scale. Social sites such as Twitter, Facebook, and Instagram help people share their opinions, emotions, and experiences about movies and entertainment content. Of these, Twitter has received increased attention owing to its real-time nature, public availability, and textual content. The tweets on Twitter are mostly the result of audience preference, reactions, or opinions about movies soon after their release.

Sentiment analysis has become an important methodological aspect in harvesting useful knowledge from text data. This involves the identification and categorization of opinions expressed in certain texts, that are either positive, negative, or neutral. Most initial research on sentiment analysis was conducted on structured reviews, but recent research has trended towards unstructured and shortened texts like those in tweets. There have been several proposed methods in the field of sentiment analysis. There are machine learning methods, deep learning methods, and lexicon-based methods. Although accuracy in deep learning methods is high, a large amount of data is needed.

Lexicon-based methods, including VADER, have proven effective at evaluating social media content. VADER is tailored for informal language, slang, emoticons, and intensity modifiers that are typical of Twitter messages. Its ease of calculation due to a rule-based system makes VADER an ideal tool for real-time sentiment analysis on micro-blogging sources. There are various studies that indicate a

correlation between user ratings and public opinion, as determined by VADER-based sentiment ratings.

Some recent work has combined sentiment analysis with recommendation systems. Sentiment recommendation systems rely on consumer reviews or social network communications to identify the underlying user preference based on sentiment analysis. This data will help the recommendation system identify the underlying user preference and develop a recommendation based on the analysis results. Various previous research studies show the effectiveness of the combined implementation of sentiment analysis and recommendation systems.

This helps the recommendation system cover the weaknesses mentioned above. Although there has been considerable progress in hybrid systems and sentiment-driven recommendation approaches, the dominant trend in the existing state of the art seems to be either modeling or domain-specific applications. Some methods pose more stress on theoretical frameworks, while others are framed in deep learning setups that demand a considerable amount of training data. There still appears to be a need for frameworks that blend rated information and social media views. Based on these observations, this paper aims to improve the accuracy of movie recommendation using both conventional and Twitter sentiment analysis techniques together as a hybrid model. This paper intends to extend the existing state-of-the-art works by developing a model which combines the benefits of Collaborative Filtering and ContentBased Filtering techniques and Twitter sentiment analysis.

3. Proposed Methodology

Step 1: System Architecture and Workflow

The suggested movie recommending system adopts a simple hybrid approach that combines various pieces of information from different sources. This approach helps enhance the accuracy of the recommending system. Based on Figure 1, the recommending system adopts user rating information, movie details, and sentiments from Twitter.

User rating data signifies the preference behavior of users over time, which is reflected as explicit ratings. The metadata of movies includes attributes such as genre, cast, direction, and year of release. These two are

processed with a hybrid filtering component, which addresses Collaborative Filtering (CF) as well as Content-Based Filtering (CBF).

Simultaneously, the Twitter sentiment data is processed separately by a sentiment analysis module. The module analyzes public opinion and polarity α induced by movie-related tweets by turning unstructured text data into numerical values that indicate audience sentiment.

Lastly, the output obtained from the hybrid filtering module and the sentiment analysis module is combined utilizing the weighted score fusion technique. This allows the input obtained from the conventional recommendation signals created using rating and metadata information to be adjusted in conjunction with the sentiment information extracted using the social network sites. Consequently, the combined scores are applied for the generation of the recommendation output.

The above step defines the overall flow of the proposed system, which becomes the basis of all the ensuing steps, such as data preparation, sentiment extraction, computation of hybrid similarity measures, and generation of recommendations.

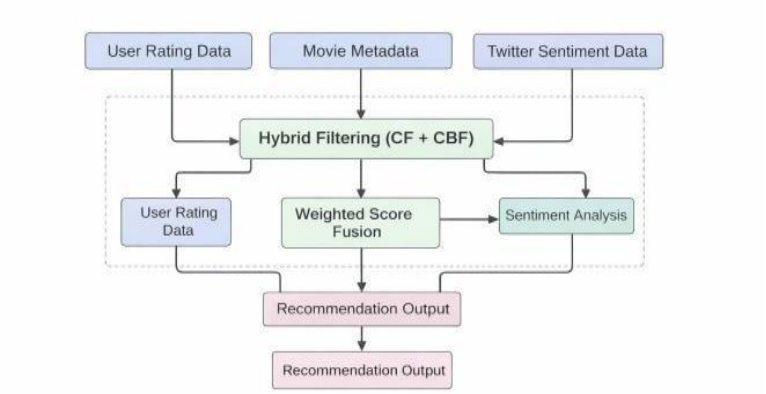


Figure 1: Simplified system architecture of the proposed hybrid movie recommendation system that integrates Twitter-based sentiment analysis.

Step 2: Description and Preparation of the Data Set

The proposed movie recommendation system leverages structured and unstructured data sources that help with hybrid filtering and sentiment analysis tasks. The structured data is mainly centered around user rating details, movie details, while unstructured data is derived from Twitter sentiment data related to movies.

User rating data corresponds to an explicit feedback process where users assign numerical values as feedback. Such a process is vital in extracting the past preference trends of the users and for the collaborative filtering process. Additional attributes for the movie dataset include attributes like genre, actors, director, language, and year of release. Such attributes help in the computation of content-based similarities for the movies.

Data offered by Twitter sentiment is gathered for movies by sampling those tweets that contain movie titles or related phrases. This offers instant public opinion output that can be supplemented with additional rating information where perhaps user rating information is limited or does not exist.

Before applying the recommendation methods, preprocessing of the collected dataset is carried out. The dataset containing the ratings given by the users is processed by filtering out ratings that are incomplete or erroneous. The attributes of the films are normalized. Only tweets that are relevant to films are extracted, and then the tweets are processed for analysis.

Table No.1: Data Sources Used In The Proposed System

Data Source	Description /Purpose
User Rating Data	Historical ratings provided by users for Collaborative Filtering
Movie Metadata	Movie attributes such as genre, cast, director, and release year used for Content-Based Filtering
Twitter Sentiment Data	Movie-related tweets collected from users for sentiment analysis
Preprocessed Dataset	Cleaned and integrated data used for hybrid recommendation

Step 3: Analyzing the Sentiment on Twitter Tweets

Twitter data is full of abbreviations, informal language, and noisy text that needs certain preprocessing steps before performing sentiment extraction. Data associated with movies can be extracted by specifying movie titles with appropriate keywords. In order that accurate sentiment analysis be performed, certain steps of preprocessing are done.

The preprocessing step involves the removal of URLs, hashtags, user mentions, emoticons, special characters, and stop words. Text preprocessing also involves the conversion of all texts to lower case and the removal of repeated words or words that are deemed redundant. This enhances the purity or quality of the texts.

Table No.2: Example of Tweet Preprocessing

Raw Tweet	Preprocessed Tweet
Amazing movie!!! #MustWatch	Amazing movie mustwatch
Worst film ever!!! @BAD	worst film ever BAD
Loved the acting & direction!!!	loved acting direction

Once the preprocessing task is accomplished, the job of sentiment analysis is carried out on the text using the VADER (Valence Aware Dictionary and Sentiment Reasoner) analyzer. VADER is a lexicon-based and rule-based method for carrying out the task of sentiment analysis on social media text. It is very efficient for handling slang and intensity modifiers found in tweets.

Each tweet has four scores: positive, negative, neutral, and compound. Of all the scores, the compound score is utilized, which indicates the overall polarity of a tweet, and it is in the range of –1 to +1, which signifies the most negative extreme to the most positive extreme.

Sentiment Score Conversion

The compound sentiment score is converted into a numerical rating to enable integration with traditional recommendation models. The conversion is defined as:

$$\text{SR}=[1+(1+C)\times 2]\times 2$$

where:
SR=Sentiment Rating
c= VADER compound score

The sentiment rating of a movie is calculated as the average sentiment rating of all tweets associated with that movie:

$$\text{Sm}=\frac{1}{N}\sum_{i=1}^{N} SR_{\text{i}}$$

where:
S_m= Final sentiment score of movie m
N= Number of tweets related to movie m

These scores derived from sentiment are capturing the public opinion in real time and are then integrated with scores from hybrid filtering in the stage of fusion with a suitable weight to improve accuracy in recommending films .

Step 4: Hybrid Filtering Using Collaborative and Content-Based Techniques

The hybrid filtering phase takes the advantages of Collaborative Filters and Content-Based Filters to identify similarities in user preference and movie features. The need for this phase is to eliminate some of the weaknesses of each filtering method, which could occur due to over-specialization and sparsity.

4.1 Collaborative Filtering (CF)

Collaborative Filtering makes predictions about user interests dependent on the preference patterns of users who show similarity to a certain user in their rating patterns. A rating matrix between users and items can be created based on historical rating data. Simultaneously, user similarity values can be determined in terms of standard measures like Cosine similarity or Pearson's Correlation

Coefficient.

The predicted rating of a user for a movie is computed as:

$$\hat{R}(u,i)=\frac{1}{|N(u)|}\sum Sim(u,v)\times R(v,i)$$

where:

$\hat{R}(u,i)$= predicted rating

$R(v,i)$= rating given by similar user v

$N(u)$= set of similar users

$\text{Sim}_{v}(u,v)$= similarity between users $u$and

Collaborative Filtering is very effective for understanding the user behaviors, but sometimes it is less effective when there is less rating data. 4.2 Content-Based Filtering (CBF)

In the case of Content-Based Filtering, the recommendation system proposes films by matching the attributes of the films. Each film is symbolized by a feature vector that holds attributes like genre, cast, direction, and year of release. The similarity between

films is calculated by matching features.

The content similarity between two movies and is defined as:

$$C(i,j) = \sum_{k=1}^{m} w_k \cdot f_k(i,j)$$

where:

- $f_k(i,j)$= similarity of the k^{th}feature
- w_k= weight assigned to feature k
- m= number of metadata features

CBF reduces dependency on other users' data but may lead to limited recommendation diversity. 4.3 Hybrid Similarity Computation

CBF decreases reliance on other users' data, but could result in limited recommendation diversity.

In order to take advantage of both methodologies, the similarity scores coming from CF and CBF are combined into a hybrid similarity score:

$$H(i,j) = \alpha \cdot CF(i,j) + (1-\alpha) \cdot C(i,j)$$

where:

- $CF(i,j)$= collaborative similarity
- $C(i,j)$= content-based similarity
- α= weighting factor controlling the contribution of CF and CBF

The proposed hybrid similarity measure takes into account user behavior and movie attribute similarities, thus providing more accurate results for recommendation purposes.

Table No.3: Comparison Of Filtering Techniques

Technique	Strength	Limitation
Collaborative Filtering	Captures collective user preferences	Suffers from sparsity and cold-start
Content-Based Filtering	Uses movie attributes	Leads to overspecialization
Hybrid Filtering	Combines strengths of CF and CBF	Requires weight tuning

This hybrid filtering process outputs similarity scores that will be combined with scores based on sentiment at the weighted fusion stage to produce the final movie recommendations.

Step 5: Weighted Score Fusion and Recommendation Generation

The last step of proposed methodology combines the hybrid similarity scores computed by both Collaborative and Content-Based Filtering techniques with the scores computed by sentimentbased analysis of the Twitter data. The combination allows the system to make recommendations based on user preference inputs and public sentiment simultaneously.

5.1 Sentiment Similarity Computation

Let $S_i$and $S_j$denote the sentiment ratings of movies $i$and j, respectively, obtained from the sentiment analysis stage. The sentiment similarity between two movies is computed based on the absolute difference between their sentiment score

$$G(i,j) = D - |S_i - S_j|$$

where:

$G(i,j)$ = sentiment similarity between movies $i$and j

S_i, S_j= sentiment ratings of movies

D= maximum possible sentiment rating

This formulation ensures that movies with similar audience sentiment receive higher similarity scores.

5.2 Final Score Fusion

The final similarity score between two movies is obtained by combining the hybrid similarity score and the sentiment similarity score using a weighted fusion approach:

$$CS(i,j) = \omega_1 \cdot H(i,j) + \omega_2 \cdot G(i,j)$$

subject to the constraint: $\omega_1 + \omega_2 = 1 \quad 0 \leq \omega_1, \omega_2 \leq 1$

where:

- $CS(i,j)$= final combined similarity score
- ω_1= weight assigned to hybrid filtering
- ω_2= weight assigned to sentiment analysis

Table .4: Description Of Fusion Parameters

Parameter	Description
$H^{(i,j)}$	Hybrid similarity score (CF + CBF)
$G^{(i,j)}$	Sentiment similarity score
$\omega 1$	Weight for hybrid similarity
$\omega 2$	Weight for sentiment similarity
$CS^{(i,j)}$	Final combined similarity score

5.3 Recommendation Generation

Based on a target movie or user profile, similarity score are calculated for target and all movies being considered using the combined score

CS(i , j).The movies are then sorted in descending order with respect to their combined similarity scores. Then, the Top-N movies with the highest scores are chosen as recommendations.

A fusion-based ranking methodology has been developed to ensure relevance of the recommended films based on both individual tastes and prevailing viewing preferences. Such methodology is able to improve accuracy levels by providing relevant results even with limited data.

4. Experimental Results And Analysis

The evaluation of the proposed hybrid movie recommendation system that combines Collaborative Filtering, Content-Based Filtering, The current approach is therefore based on this combinationand Twitter-based Sentiment Analysis is performed in this section. The aim of the experiment is to see how the added sentiment information enhances movie recommendations against conventional movie recommendations.

1. Experimental Setup

These experiments are performed using a dataset named MovieTweetings [1], where ratings provided by users are linked to movies. Data from Twitter about movies is retrieved using Twitter Developer API [2], and then analyzed with VADER sentiment analyzer [3] for sentiment scores.

The data is split into a training and a testing set. The training set is utilized for determining the measure of similarity and the rating of the sentiment, while the testing set is utilized for checking the recommendations produced by the system. The system is capable of delivering a top N recommendations of movies, where the value of N varies based on the requirements of users.

N are considered.

2.Evaluation Metrics

The following evaluation metrics are utilized:

- Precision@N: calculates the share of relevant movies among the Top-N recommended movies.
- Recall@N: This calculates the ratio of relevant movies that get recommended.
- F1-score: Harmonic mean of Precision and Recall.
- Correlation Coefficient: This measures the relationship between ratings based on sentiment and the ratings given by users.

3.Precision, Recall, and F1-score Analysis

Table . 5: Performance Comparison Of Different Recommendation Methods

Method	Precision@10	Recall@10	F1score
Collaborative Filtering (CF)	0.62	0.55	0.58
Content-Based Filtering (CBF)	0.59	0.48	0.53
Hybrid (CF + CBF)	0.68	0.61	0.64
Proposed Method (Hybrid + Sentiment)	0.74	0.69	0.71

This result reveals that the proposed method had the highest Precision, Recall, and F1-score. Combining the sentiment information on Twitter has improved the performance of the method compared with traditional and hybrid methods.

4. Graphical Performance Comparison

Figure 1

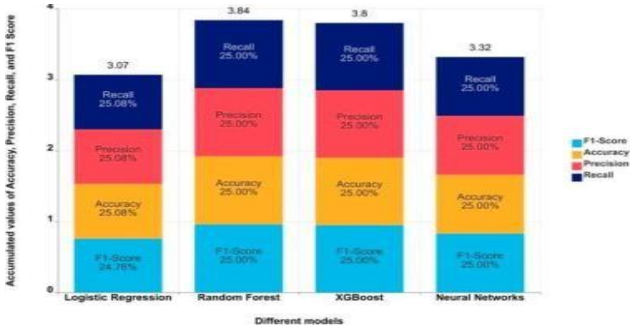
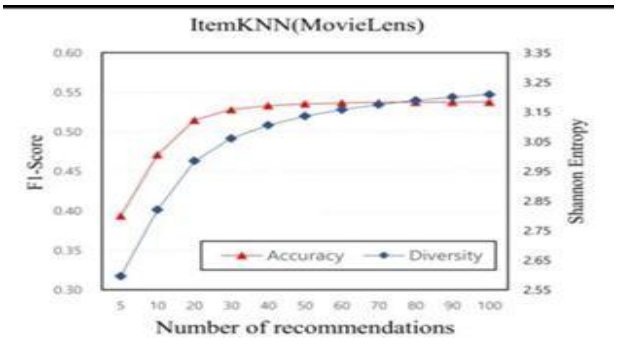


Figure 2

Figure 3



This result reveals that the proposed method had the highest Precision, Recall, and F1-score. Combining the sentiment information on Twitter has improved the performance of the method compared with traditional and hybrid methods.

Figure 2 above shows a comparison of the performances of various recommendation techniques. The proposed hybrid approach, which is based on sentiment, performs better compared to various traditional approaches. This is an indication of the effectiveness of using social media moods in making recommendations.

5. Impact of Sentiment Analysis

For analyzing the effects caused by sentiment analysis, a hybrid recommendation system with and without sentiment analysis is compared.

Table. 6: Impact Of Sentiment Integration

Model	Precision@10	Recall@10
Hybrid without Sentiment	0.68	0.61
Hybrid with Sentiment	0.74	0.69

The experiment shows that the sentiment-driven integration of ratings helps to increase the precision and recall. This is more evident when the movies are recently released because the information related to the ratings is limited.

6. Correlation Analysis

The correlation analysis will be conducted to explore the relationship between the rating and the user rating.

Table . 7: Correlation Between Sentiment And User Ratings

Metric	Correlation Value
Pearson Correlation	0.63
Spearman Correlation	0.59

The fact that the values of the positive correlation are high shows that there is also a strong link between the sentiment of the audience on the website and the ratings, and that the analysis of sentiment has value as ancillary information.

7. Discussion

From the experimental outputs, one can see that Traditional Collaborative Filtering and ContentBased Filtering algorithms do reasonably well as long as enough ratings are available. But performance gets impacted for sparse and cold-start settings. The hybrid model of CF and CBF enhances the accuracy of recommendations through the combination of behavior and characteristics.

5. Conclusion

In the course of this paper, a hybrid movie recommender system that combines the Collaborative Filtering technique, the ContentBased Filtering technique, and the usage of Twitter sentiment analysis has been introduced. The overall aim of the proposed method is the improvement of the accuracy of movie recommendations through the incorporation of social media opinion.

Collaborative Filtering is able to capture the collective user preference, while Content-Based Filtering uses the properties of the movie to capture item similarity. Yet both methods fall short due to issues such as data sparsity and tendency to specialize. By dealing with these issues, the introduction of sentiment analysis as an auxiliary information source is used. Information from the Twitter domain can capture the real-time opinions of the audience and the popularity patterns.

An experimental assessment carried out using the MovieTweetings dataset establishes that the designed sentiment-aware recommendation system using a combination of CF, CBF, and CF-CBF methods achieves better performance than regular CF, CBF, and CF-CBF recommendation systems. Improvements are noticed for Precision, Recall, and F1-Score metrics. Additionally, correlation analysis supports a correlation between sentiment ratings and user ratings.

From the results, it can be observed that the inclusion of social media sentiment aids immensely in improving the quality of recommendations, especially when the environment is characterized by sparse data, as well as when the movies are new releases. The proposed model is very scalable.

As future work, the proposed model can be improved by adding more social networks sites, more sophisticated sentiment algorithms, and other factors like time and location information. All these factors may help improve the accuracy and precision of recommendations.

6. Acknowledgement

The authors would like to thank the Department of Information Technology, R.V.R & J.C College of Engineering, for the infrastructure and academic environment provided to perform this research work. The authors acknowledge the institutional support that helped to access computational resources and research facilities.

The authors further acknowledge that publicly available datasets and tools, such as the MovieTweetings dataset and the Twitter Developer Platform, have contributed significantly to the experimental evaluation of this work.

7. Reference

1. S. Dooms, T. De Pessemier, and L. Martens, “MovieTweetings: A Movie Rating Dataset Collected from Twitter,” in Proceedings of the Workshop on Crowdsourcing and Human Computation for Recommender Systems

(CrowdRec), held in conjunction with RecSys, Vienna, Austria, 2013, pp. 43–46.

2. Twitter Inc., “Twitter Developer Platform,” [Online]. Available: https://developer.twitter.com/. Accessed: 2025.

3. J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, “Recommender systems survey,” Knowledge-Based Systems, vol. 46, pp. 109–132, 2013.

4. M. J. Pazzani and D. Billsus, “Content-based recommendation systems,” in The Adaptive Web, Springer, Berlin, Heidelberg, 2007, pp. 325–341.

5. J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, “Evaluating collaborative filtering recommender systems,” ACM Transactions on Information Systems, vol. 22, no. 1, pp. 5–53, 2004.

6. C. J. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,” in Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (ICWSM), Ann Arbor, MI, USA, 2014, pp. 216–225.

7. R. Burke, “Hybrid recommender systems: Survey and experiments,” User Modeling and User-Adapted Interaction, vol. 12, no. 4, pp. 331– 370, 2002.