MOVIE SUGGESTED PLATFORM USING SENTIMENT ANALYSIS FROM MICROBLOGING

M.Pradeep Kumar[1]., P.Ramya[2].,B.Sneha[3].,P.Sri Rekha[4].,G.Roshini[5]

1 Assistant Professor, Department Of CSE., Malla Reddy College Of Engineering For Women., Maisammaguda.,

Medchal.,Ts, India (pradeepit09@gmail.com)

2, 3, 4, 5 B.Tech CSE, (19RG1A05B0, 19RG1A0565, 19RG1A05A8, 19RG1A0575),

Malla Reddy College Of Engineering For Women., Maisammaguda., Medchal., Ts, India

Abstract

Intelligent recommendation systems play a crucial role in meeting consumers' needs for curated content. Collaborative filtering and content-based filtering are two of the most conventional methods used in recommendation systems. However, these methods have some restrictions, such as the requirement of user history and habits to accomplish the task of recommendation. This research presents a hybrid recommendation system that incorporates collaborative filtering, content-based filtering, and sentiment analysis of movie tweets to lessen the impact of such dependencies. In order to gauge audience reaction and track industry trends, we culled tweets about the film from several microblogging services. Experiments using public databases have shown encouraging results.

One, Preface

The internet is become an integral component of modern society. The abundance of information accessible to users is a common complaint. To aid end users in making sense of all this data, recommendation systems (RS) are being rolled out. The most common uses for RS are in information management and e-commerce platforms, including travel and entertainment sites. Emphasis this article, we focus emphasis on RS since movies provide so much needed downtime and amusement. Online movie recommendation systems are indispensable. Genres including humour, suspense, animation, and action help make movies distinct from one another. Metadata, such the film's release year, language, director, and cast, may also be used as a means of organization. Most video-streaming websites offer users a selection of films that are similar to those they have already watched or rated highly. Movie recommendation systems [23, 25, 48, 3, and 53] make it easier to choose entertaining films without having to go through many options.

Providing users with recommendations for films that would likely appeal to them is the most important function of a movie recommendation system. With the exponential growth of available web data, RS have become more useful for making judgments in many areas of daily life. Content-based filtering (CBF) and Collaborative filtering (CF) are the two main types of RS.

The notion of CF originated from the proclivity of humans to base their decision-making on empirical evidence, established norms, and widely disseminated knowledge found on the internet. The notion of CF was first proposed in Netnews by Resnick et al. [37], with the aim of assisting readers in locating items of interest among the vast quantity accessible. CF aid decision-making by incorporating the experiences and perspectives of others. In CBF [46], things are recommended based on the similarity among the contentual information of the objects, however when two users have comparable ratings for items, they are regarded to be like-minded. People's day-to-day emotional states are increasingly being documented online because to the proliferation of social media sites like Quora, Facebook, and Twitter. Twitter [1, 2, 13] was launched in 2006 and has since become one of the most popular social media platforms due to its 140-character restriction on user-generated content. Twitter's USP is that its users may tap into a wealth of user-generated content in addition to receiving updates based on their social connections. Twitter's users stay in the know about their favorite personalities, shows, and news with the use of short messages called tweets. In this research, we offer a movie recommendation method that combines hybrid scores with sentiment analysis of tweets about movies pulled from the MovieTweetings database.

The following are the paper's most significant contributions:

By fusing collaborative filtering with content-based filtering, we present a novel hybrid recommendation system.

Second, this recommendation system is improved with the use of sentiment analysis.

Third, substantial experimental analysis is given to shed light on the suggested recommendation system. Last but not least, a quantitative and qualitative comparison to other baseline models is presented.

Following this structure, the rest of the paper will discuss: The relevant research was summarized in Section 2. Section 3 lays out the methods that will be used. In Section 4, we provide the outcomes of using the suggested paradigm. In Section 5, the analysis is summed up.

2 Related Papers

The best knowledge management systems are recommender systems, which analyze a user's browsing history and other data to provide suggestions tailored to that individual. In recent decades, several RSs have been introduced. CF, CBF, hybrid, and sentiment analysis systems were among the methods used by these tools. The next section elaborates on this topic.

2.1 Types of Filtering: Collaborative, Content-Based, and Hybrid

Several RS strategies for product suggestions have been proposed in the literature [41]. The first documented usage of collaborative filtering was in [16], which suggested a search engine dependent on document content and user feedback. Combining pre- and post-purchase ratings into a single RS for e-commerce applications was proposed by Guo et al. [18]. Soleymani et al. [43] developed an emotional rating of movie sequences that takes into account both the user's emotional state and the video's content. They discovered that the users' self-reported levels of arousal and valence were associated with the movie sequences. In order to classify and score video information, physiological data from the periphery were used.

Implicit evaluations were deduced by Yang et al. [51] from users' page views.

Users are thought to enjoy papers better if they read more of them. This idea is useful for combating CF's cold start issue. An RS optimization issue is not well-posed. Gray wolf optimization [22], artificial bee colony [19], particle swarm optimization [45], and genetic algorithms [6] are only a few of the optimization techniques that have been developed by researchers. Using the bio-inspired gray wolf optimizer and fuzzy c-mean (FCM) clustering algorithms, Katarya et al. [22] created a movie recommendation system based on collaborative filtering. The initial cluster location was determined using the Gray wolf optimizer. Predictions of users' evaluations for a given film were further informed by their past behavior and the similarities between them. In [21], the authors propose an enhanced architecture using ABC-KM (artificial bee colony and k-mean cluster) for a collaborative movie selection system, simplifying scalability and cold start issues. When compared to preexisting frameworks, the hybrid cluster and optimization technique combination demonstrated improved movie prediction accuracy.

One of the most popular and extensively studied recommendation system paradigms is content-based filtering [49, 27, 33, 47]. This RS method requires a detailed product description in addition to a user preference profile. The ability of words to discriminate is examined by the authors in [31]. In order to retrieve relevant materials, they use a weighting method based on the titles, abstracts, and full texts to get to this conclusion. To generate music suggestions, Cantador et al. [9] employ user and item profiles expressed as weighted collections of social tags. Articles for home renovation were suggested using a Personalized Recommender System (PRES) created by Meteran et al. [46], which used a combination of the TF-IDF and the cosine similarity to estimate the degree to which a user profile vector and a document were comparable. A novel technique for news item recommendation was developed by Goossen et al. [17], called CF-IDF, and it makes use of TF-IDF and a domain ontology. When tested, assessed, and implemented on the Athena framework, the performance of this technique surpassed that of the TF-IDF approach on a variety of metrics, including accuracy, recall, and the F1-measures.

Recent studies [34, 38, 44, 5, 7] have shown that a hybrid strategy is superior than more conventional methods. Hybrid systems' primary benefit is that they use a combination of several recommendation methods, each of which has its own limitations. The authors of [30] created a content-boosted CF System that relied only on content-based features inside a cooperative architecture. This method enhanced previous enhancements to prediction, first-rater, and sparsity. The user's input is used to generate N recommendations, and the user's preferences are tracked until no recommendation is selected. This system was created by Zhang et al. [52]. The mobile recommender system described in [32] blends a hybrid recommendation engine with a mobile 3D geographic information system architecture.

Twenty-seven users, spanning the ages 24-48, were chosen to put the suggested framework through its paces. Users were given tasks along a highway to complete, such as locating a hotel, bar, or restaurant, in order to gauge the effectiveness of the system. According to user reviews, the 3D map-based interface works well and is able to function well despite the small screen size of most mobile devices.

2.2 Emotional evaluation

The analysis of sentiment [35, 8, 29, 36, 15, 50] is a method often employed by academics to gather the views of the general public. Online software services have been rated using sentiment analysis [14]. By using objective metrics like sentiment analysis and subjective reasoning, their work improves CBF and CF algorithms. The polarity and certainty of review sentences are determined using a sentiment analysis algorithm [28]. In [20], the authors suggest a sentiment analysis approach called the Valence Aware Dictionary and Sentiment Reasoner (VADER). Five overarching principles were derived from lexical components; these rules represent established grammatical and syntactic norms for conveying and emphasizing strong emotions. Classifying the tweets as good, neutral, or negative yielded an F1 score of 0.96. Based on Twitter data, the author of [4] suggests an automated feedback mechanism.

Twitter comments were classified using a variety of methods, including the Support Vector Machine, Naive Bayes, and Maximum Entropy. Based on the user's current emotional state, the authors of [39] presented a music recommendation system for mobile devices. Two hundred people (hundred men and women) were surveyed after they selected their preferred musical genre for their online profiles. In the end, the profile data showed that 91% of the participants were satisfied with the experience. The cold start issue in a CF system is addressed in [26], where the author proposes the KBridge architecture. In this framework, microblogging posts were also subjected to sentiment analysis, with a 1–5 polarity score being assigned to each. Connecting users' social media and television viewing habits may lead to a more effective recommendation system, as shown by the study's findings. To facilitate the incorporation of sentiment analysis and CF, the author of [24] presented a rating inference technique to convert textual reviews into ratings.

Our suggested model is a hybrid recommender system where sentiment analysis score is used to further improve recommendations. The reliability and efficiency of our technique have been confirmed by both quantitative and qualitative experimental assessments.

3 Methods Put Forth

Figure 1 depicts the suggested sentiment-based recommendation system. Here, we lay out the methodology behind the recommended system and its constituent parts.

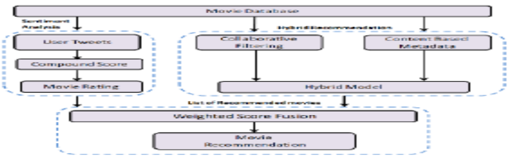

Fig. 1: Proposed movie recommendation framework.

3.1 Explanation of the Dataset

The suggested system requires two types of databases. One is Twitter user-tweets, and another is a database of user-rated movies that includes ratings for all relevant movies.

The 3.1.1 Open Registry

In the field of movie recommendations, numerous prominent public databases have seen extensive usage. The suggested framework includes a method for retrieving tweets about movies that already exist in the database. Tweets about recent films were utilized to include sentiment analysis into the suggested recommendation system.

Our research was not supported by experiments performed on publicly available datasets such the Movielens 100K 1, Movielens 20M 2, the Internet Movie Database (IMDb3), or the Netflix database 4. These archives comprised classic films for which there was no relevant microblogging data. The MovieTweetings database [12] was chosen for the proposed solution after extensive evaluation of other databases.

The MovieTweetings database is often compared to the more traditional MovieLens.

In contrast to MovieLens, where at least 20 films have been reviewed by a single user, MovieTweetings is an unfiltered database. In order to give more accurate data for sentiment analysis, this database strives to include the most recent ratings. The data set comes straight from the world of social media. It's quite varied yet rather sparse. Information about the MovieTweetings database is shown in Table 1.

Table 1: Details of MovieTweetings database.

Metric	Value
Ratings	646410
Unique Users	51081
Unique Movies	29228
Start Year	1894
End Year	2017

3.1.2 Modified MovieTweetings database

Our suggested project involves modifying the MovieTweetings database in order to construct the recommendation system. The change was prompted mostly by the need to include user-tweet sentiment analysis into the prediction of movie recommendations. Films released between 1894 and 2017 are included in the MovieTweetings archive. Due to the dearth of tweets for older films, we restricted our analysis to those that were released in 2014 or later, and we only used a portion of the MovieTweetings database that met our criteria.

$$release\_year_{movies} \geq 2014 \quad (1)$$

Following this database manipulation, a new, more suitable database for the recommendation system's implementation was created. There were 6209 movies total with ratings from 292863 individuals in this updated database. Movies released in the years 2014–2017 account for 45% of all ratings, which is around 20% of all movies in the MovieTweetings database. No earlier or later films were considered for inclusion in this study. Movies are shown in Fig. 2 according to the years they were released. There are three distinct sections to the MovieTweetings database. The user-to-Twitter-id mapping is the first piece of the puzzle. The user ratings for individual movie elements make up the second part, while data on the rated films themselves make up the third.

Fig. 2: Year wise movies distribution.

The information scraped from MovieTweetings includes user reviews of movies and their classifications. However, aside from the year of release and the genres covered, it provides no additional information. Only in the scenario of social filtering, when there are many users with similar traits, is such information effective. In the case of collaborative filtering, in which recommendations are based on the perceived similarities between users rather than on the ratings of any one user, this information might be quite useful. The similarity of movies based on their qualities and the socially filtered data have both been incorporated in the suggested approach. Movie data was retrieved via the Movie Database's (TMDb) application programming interface. TMDb 5 is the go-to place for in-depth movie information, and it supports over 30 different languages!

Table 2 displays the MovieTweetings database's film characteristics.

Extremely niche films from a variety of countries and languages, for which metadata is not present in TMDb, have been added to the modified database. Videos with practically little information were deleted. About 4,500 films made it into the final database.

Information about the database's statistics may be seen in Table 3.

Table 2: Example of a movie entry in the modified MovieTweetings database.

Attribute	Value
MovieID	0451279
Title	Wonder Woman
Runtime	141 min
Genre	Action,Adventure,Fantasy
Director	Patty Jenkins
Writer	Allan Heinberg
Actors	Gal Gadot,Chris Pine
Rating	7.6 Massachusetts Institute of Technology in 1996.
Production Companies	DC Films,Tencent Pictures
Popularity	524.772
Language	en
Production Countries	United States of America
Budget	816303142

Table 3: Details of the modified MovieTweetings database.

Metric	Value
Ratings	292863
Unique Users	51081
Unique Movies	4515
Start Year	2014
End Year	2017

3.2 Analysis of user tweets

As can be seen in Fig. 3, we used Twitter API 6 to get tweets about movies from MovieTweetings. Huge quantities of noise, including hashtags, emoji, repeating phrases, and other unnecessary data, were eliminated from the tweets before they were processed.

3.2.1 Tweets are processed beforehand

Inaccurate sentiment analysis was hindered by the abundance of irrelevant and distracting details present in tweets. Common examples are filler words, commas, periods, hyperlinks, accents, and repetitive phrases. Table 4 provides a few instances of distracting and irrelevant material. The tweets' collected text underwent sentiment analysis after preprocessing, and the results were included into the recommendation system's development.

Table 4: Examples of noisy and uninformative parts in tweets.

Types of noise	Example
Stop words	a, and, the, after, am
Lemma	serve, served and serving
Web links	www.tripadvisor.com
Filtering of repeating words	happyyyy, heloooo
Special Characters	!, @, #, $, %, and _

3.2.2 Sentiment analysis of user tweets

VADER is a rule-based technique and vocabulary for sentiment analysis that is both computationally efficient and useful. In this method, the sentiment analysis will take care of identifying tweets with a subjective tone. VADER creates an emotional lexicon or dictionary that may be used to link words to emotions. The dictionary analyzes the tone of words and tweets. Either the sentiment class or the tweet score may be analyzed using lexical methods. VADER extracts four emotional facets from each tweet. The first three of these categories are good, neutral, and negative, respectively. The fourth factor is a compound score, defined as the mean, median, and mode of the first three metrics. The composite score may be anywhere from -1 (the least desired) to +2 (the most desired) for each movie. Then, using Eq. (2), the score is transformed into a numeric value between 1 and 10.

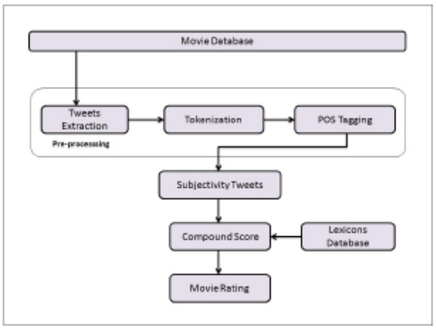

Fig. 3: A framework of the VADER sentiment analysis system.

$$Rating = [1 + (1 + x) \times 2] \times 2 \qquad (2)$$

If x is a composite score, then...

In addition, VADER runs quicker than alternatives like the Naive Bayes analyzer and TextBlob. The use of slang, emoticons, and acronyms in writing is where they really shine.

3.3 Suggestion for a Hybrid

Here, we detail how a hybrid recommendation model was developed by fusing content-based similarity characteristics with collaborative social filtering. Let

$$f = \{f_1, f_2, \cdots, f_n\} \text{ and } q = \{q_1, q_2, \cdots, q_n\}$$

are feature vectors depending on the content and the weight vector.

To measure how similar two things i and j are, we use the formula:

$$C(i,j) = \begin{cases} \sum_{n=1}^{N} f_n(A_{n_i}, A_{n_j}), & \text{for } i \neq j \\ 0, & otherwise \end{cases} \qquad (3)$$

where fn(Ani ;Anj) represents the similarity between two items' feature values Ani and Anj. Metadata or other item-specific information is used to calculate distance in Eq. (3). Fij are built by summing the item-level proximity vector C and multiplying by the item-level weight vectors q. Where n and M are the number of feature attributes and the number of items, respectively, Fij is a feature matrix with dimensions n _ M(M1) 2. Using a user-generated social graph, the weight vectors q are analyzed to see how well they represent the opinions of the users. If ui is a user in the database, then U = fu1; u2; _ _ _ ; ung. For M items, we build a user-item matrix. The user-item matrix has a high degree of sparsity, which is an essential attribute. This User-Item matrix is used by standard collaborative Filtering [42] to estimate a user's rating of an item i based on the ratings of the users in the user's neighborhood, typically K people. Cosine similarity, Pearson correlation, and other similarity measurements are used to identify nearby users. The aggregated ratings after picking K nearby people are weighted as follows:

$$rating(u,i) = \frac{1}{K}\sum_{k\in K} similarity(user_u, user_{v_k}) \cdot rating_{v_{ki}} \qquad (4)$$

where u and vk denote the intended audience and the K closest friends, respectively. To avoid the sparsity of the User-Item matrix while making rating predictions, we use the method of collaborative filtering. We use the modified User- Item matrix to build a social network using products as nodes. The degree to which users find two objects to be same is graphically shown here. The social network is followed when deciding how much weight to give to each attribute.

We create a framework as shown in Eq. (5) to get the best feature weights q:

$$S(i,j) = q \cdot F_{ij} \qquad (5)$$

which can be expanded as:

$$S(i,j) = q_1 \cdot f_1(A_{1i}, A_{1j}) + q_2 \cdot f_2(A_{2i}, A_{2j}) + \cdots + q_n \cdot f_n(A_{ni}, A_{nj}) \qquad (6)$$

Now we can assess the User-Item matrix for a pair of items i and j, where S(i; j) is the number of people that are interested in both i and j. Dimensionally, S is a matrix with the form 1 _ M(M1) 2, q is a matrix with the form 1 _ n, where n is the number of content-based features, and F is a matrix with the form 1 _ M(M1).

$$n \times \frac{M(M-1)}{2}.$$

Using the Moore-Penrose Pseudoinverse as shown in Eq. (7), we get the weight vector q for all the metadata feature characteristics for the whole items:

$$q = S^{-1} \cdot F \qquad (7)$$

3.4 Weighted score fusion

In section 3.3, we discussed how we used the movie-similarity and user-similarity paradigms to calculate the feature vector weights q. The suggested system employs the idea of sentiment-fusion, with weights q standardized between [0, 1]. A sentiment rating is made up for all of M movies based on the tweets that were gathered from users. If we replace si with the movie i rating obtained from Eq. (2), we get S 2 fs1; s2; _ _ _ ; sng. According to Eq. (8), we may create a function G(i; j) between two movies i, j based on their respective sentiment scores si and sj.

$$G(i,j) = D - |s_i - s_j| \qquad (8)$$

in which D is a fixed value. Due to the 1–10 rating range, we pick 10 for the constant D in Eq. (8). Then there's the other function H(i, j):

$$H(i,j) = q \cdot f_{ij} \qquad (9)$$

where fij represents the feature similarity between movies i and j, and j and q represent the ideal weights as calculated by Eq. (7). In Eq. (10), we can see how the ultimate combined similarity CS(i; j) is defined. Combining the predefined functions G and H with appropriate weights.

$$CS(i,j) = \omega_1 \cdot H(i,j) + \omega_2 \cdot G(i,j) \tag{10}$$

$$\omega_1 + \omega_2 = 1, \qquad \omega_1, \omega_2 \in [0,1] \tag{11}$$

where!1 represents the importance of the hybrid model's similarity score and!2 represents the importance of the emotion similarity score.

4 The Analysis of Experimental Findings

Various results for the correlation coefficient between sentiment ratings and IMDb movie ratings are presented here. We also detail a complete quantitative and qualitative study that we conducted on a movie database to demonstrate the usefulness and accuracy of our suggested methodology.

4.1 The Feeling of Approval and IMDb Ratings

We ran the numbers to see whether there was a connection between how people felt about a film and how well it was received statistically. Values of the correlation coefficient may swing from minus one to plus one. Let's say we have two variables: the number of films in the database (N) and the database (D). Spearman's rank order correlation coefficient (SROCC), Kendall's rank correlation coefficient (KRCC), and Pearson's linear correlation coefficient (PLCC) are the three statistical measures of correlation. The values of the various co-relation coefficients we used are shown in Table 5. Our research indicates a beneficial relationship between subjective feelings and evaluations of films. xi and yi are the PLCC sentiment score and IMDb movie rating for the ith film, respectively, with _x and _y being the mean sentiment score and mean film rating, respectively. The ith movie's SROCC sentiment rating minus its movie rating is denoted as di. Number of concordant pairings (Nc) and number of discordant pairs (Nd) in the database are used to calculate KRCC. Movie reviews and audience ratings are correlated in Table 5.

Correlation coefficient	Definition	Value
PLCC	$\frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}}$	0.76
SROCC	$1-\frac{6}{N(N^2-1)}\sum_{i=1}^{N} d_i^2$	0.72
KRCC	$\frac{2(N_c-N_d)}{N(N-1)}$	0.51

4.2 Evaluation metric

Instead of explicitly forecasting rating levels, the system suggests applicable suggestions in many real-world situations. Top-N recommendations [11, 10, 40] propose particular goods to consumers based on their preferences. RMSE (Root Mean Squared Error) and MSE (Mean Squared Error) are often utilized in much of the mainstream literature. However, Top-N recommendation systems care less about these kinds of error metrics. Consequently, evaluation metrics (like precision) make use of the direct alternative methodologies. Accuracy is measured by the consistency between the movies you want to see and the movies the model suggests you watch. Let's call the movies that are really worth seeing (Lrel) and the ones that aren't (Lrec). Precision@N is specified in the proposed system as follows in Eq. (12):

$$Precision@N = \frac{L_{rel} \cap L_{rec}}{L_{rec}} \tag{12}$$

The Precision@5 and Precision@10 for a variety of weight combinations inform the choice of weights in the fusion in Eq. (10).Following Equation (11), we have.1 and!2.

How to Choose Fusion Weights (4.3)

The Top-N list of recommended films is ranked for each film using Eq. (10). Experiments on the metric discussed in Section 4 determine the values of!1 and!2 in Eq. (10).Using Eq. (12), we can calculate the Precision@N. All movie suggestions are culled from public sources like IMDb and TMDb. The movies that are endorsed by these two organizations are taken as gospel.

We evaluate the differences between the Precision@5 and Precision@10 over a range of!1 and!2 values. For!1 and!2, we choose the ones where the accuracy is highest.

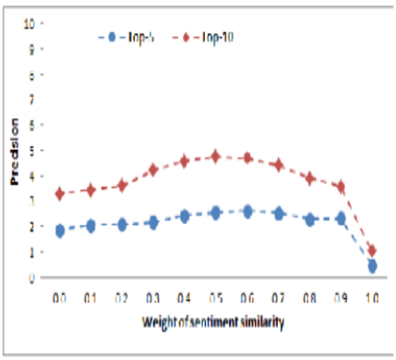

Fig. 4: Precision of Top-5 and Top-10 movies with varying sentiment similarity weights.

Weight value accuracy is maximized between 0:5 and 0:6, as can be seen in Fig. 4. Therefore, in the proposed system, we choose!1 and!2 to be 0.5.

4.4 A Look at the Numbers

Our suggested system is compared to existing models in this section. The Pure Hybrid Model and the Sentiment Similarity model serve as the baselines for our investigations. The Pure Hybrid Model (PH Model) combines content-based filtering with collaborative filtering, with the two filtering techniques reinforcing one other to produce a more reliable whole. Movies are suggested based on shared characteristics, such as genre, director, actor, and so on. As mentioned in Section 3.3, we use a social graph to assign values to the similarities. The Sentiment Similarity model (SS Model) is a movie recommendation system that ranks films purely on how similarly their tweets about those films are to one another. Precision@5 and Precision@10 are used to assess the accuracy of our suggested approach. Our suggested system's quantitative compared findings to the baseline models are shown in Fig. 5. Both the SS Model and the PH Model have average accuracy values of 0.54 and 1.86 for accuracy@5, respectively. The average Precision@10 value for SS Models is 1.04, whereas the average Precision@10 value for PH Model is 3.31. When compared to the PH and SS Models, our proposed model achieves a higher precision value in both cases, 2.54 for the Top-5 and 4.97 for the Top-10. Based on the broad consensus, as verified by the IMDb and TMDb databases, we may conclude that our algorithm will recommend at least 2 out of 5 and 5 out of 10 movies.

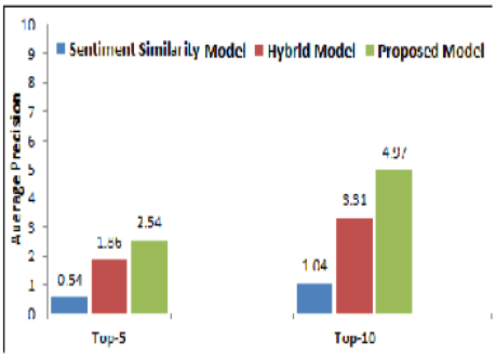

Fig. 5: Comparison of proposed model with baselines models.

4.5 Qualitative analysis

Here, we present some qualitative results from the proposed system's movie recommendations. Tables 6 and 7 demonstrate that there are foreign language films included as well. These tables show that the proposed system's movie suggestions overlap significantly with those of both IMDb and TMDb. Qualitative study of both Hollywood and Bollywood (Indian) films demonstrates the capacity to provide suggestions in a variety of contexts.

5 Last thoughts and further development

In the present day, when there is an abundance of data, recommender systems play a crucial role as a filtering medium. This case

Table 6: Qualitative analysis of Wonder Woman movie: Language (English). Movies in bold are intersecting with either IMDb or TMDb

IMDb	TMDb	Recommendations from the proposed system
Justice League	**Guardians of the Galaxy Vol. 2**	**Batman v Superman: Dawn of Justice**
Batman v Superman: Dawn of Justice	Spider-Man: Homecoming	**Suicide Squad**
Suicide Squad	Logan	**Thor: Ragnarok**
Thor: Ragnarok	**Thor: Ragnarok**	**Justice League**
Spider-Man: Homecoming	**Justice League**	Warcraft
Deadpool	Pirates of the Caribbean: Dead Men Tell No Tales	**Doctor Strange**
Logan	**Doctor Strange**	**Guardians of the Galaxy Vol. 2**
Captain America: Civil War	Baby Driver	**Kong: Skull Island**
Doctor Strange	**Kong: Skull Island**	The LEGO Batman Movie
Guardians of the Galaxy Vol. 2	Life	Batman and Harley Quinn

Table 7: Qualitative analysis of Neerja movie: Language (Hindi). Movies in bold are intersecting with either IMDb or TMDb

IMDb	TMDb	Recommendations from the proposed system
Airlift	**Airlift**	**Simran**
Pink	Pink	**Fan**
Kapoor & Sons	Rustom	**Raabta**
Udta Punjab	Ghayal Once Again	**Udta Punjab**
Drishyam	Mary Kom	**Rocky Handsome**
Rustom	**Udta Punjab**	**Rangoon**
M.S. Dhoni: The Untold Story	**Force 2**	**Raabta**
Raabta	**Fan**	**Force 2**
Dear Zindagi	**Rocky Handsome**	Te3n
Rangoon	**Simran**	**Airlift**

In this research, we offer a recommender system for movies that combines Twitter sentiment analysis data with movie information and a social graph.

The results of a sentiment analysis on audience reactions to a film have been found to be informative. Weighted score fusion was employed by the suggested system to better recommend items. The average Top-5 and Top-10 accuracy for sentiment similarity, the hybrid model, and the suggested model are.54 and 1.04, respectively, based on our studies. We compared the suggested model to many others and found that it made more accurate recommendations. To further enhance the recommendation system, we want to collect additional information about the user's emotional tone from other social media sites in the future and analyze it using sentiment analysis. This work demonstrates that there is room for investigation into users' emotional states and the incorporation of this data into e-commerce recommendation systems. Real-time updates to social graph-based weights are another area of study. In our studies, we have employed a database that does not change, and we have only analyzed films from 2017 and after.

This structure can be investigated within a shifting paradigm where the routine incorporation of newer films is possible.

Observance of moral requirements

Interests that are at odds The authors have stated that they have no financial or personal stake in the outcome of this study.

References

One each: Fabian Abel, Qi Gao, Geert-Jan Houben, and Ke Tao. Twitter user modeling for news suggestions analysis. User Modeling, Adaptation, and Personalization: Proceedings of the International Conference, pages 1-12. 2011 Springer.

Ke Tao, Geert-Jan Houben, Qi Gao, and Fabian Abel are the other two. Modeling Twitter users to propose stories to them. Pages 2962–2966, 2013 International Joint Conference on Artificial Intelligence.

Gediminas Adomavicius, Alexander Tuzhilin, and the third author. A look at the current status of recommender systems and where they may go in the future.

Published in 2005 at 17(6):734-749 in IEEE Transactions on Knowledge and Data Engineering.

Aswathy Wilson, Vinay James, Nivya Johny, Minara P Anto, and KM Muhsina. Evaluation of products based on customer feedback. Pages 3458–3462 in Proceedings of the International Conference on Electrical, Electronics, and Optimization Techniques. IEEE, 2016.

There are five of them: Ehsan Aslanian, Mohammadreza Radmanesh, Mahdi Jalili, and the fifth one is Mahdi Jalili. Relational hybrid recommender systems for content. Industrial Informatics: 2016 IEEE Transactions on.

Javier Alcal'a, Antonio Hernando, Jesus Bobadilla, and Fernando Ortega make up the rest of the gang of six. Using evolutionary algorithms to enhance the performance of collaborative filtering recommender systems. 2011;24(8):1310-1316, Knowledge-Based Systems.

Hybrid recommender systems: A survey and experimental results 7 by Robin Burke. 12(4):331-370, 2002, User modeling and user-adapted interaction.

8th place: Erik Cambria. Sentiment analysis and emotional computing. Intelligent Systems, Volume 31(2), Issue 2, 2016, IEEE.

Ivan Cantador, Alejandro Bellogn, and David Vallet make up our tenth spot. In social tagging systems, content-based suggestions are made. Specifically, on pages 237–240 of the Recommender Systems: Fourth International Conference Proceedings. ACM, 2010.

Ten. Roberto Turrin, Paolo Cremonesi, and Yehuda Koren. Analysis of recommender algorithms' efficiency in top-n recommendation problems. Pages 39–46 in the Recommender Systems 4th Conference Proceedings. ISBN 978-1-60558-906-0, ACM, 2010.

Eleventh, George Karypis and Mukund Deshpande. Recommendation algorithms based on a top n list of items. ISSN 1046-8188, 22(1):143-177, 2004. ACM Transactions on Information Systems.

Toon De Pessemier, Luc Martens, and Simon Dooms make up the group of 12. Movietweetings is a Twitter-based movie review dataset. A 2013 Workshop on Human Computation and Crowdsourcing in Recommender Systems.

Xin Fu and Yun Shen, number 13. Fuzzy method research on Twitter group behavior. 2014, Volume 25 Issue 7-8, Pages 1603-1614 of Neural Computing and Applications.

Number Fourteen: Lahiru S. Gallege and Rajeev R. Raje. An approach to choosing and recommending cloud-based software services based on objective criteria. On page 23 of the 2016 proceedings of the Eleventh Annual Cyber and Information Security Research Conference.

Balasubramanian Raman, Debi Prosad Dogra, Priyanka Singh, Himaanshu Gauba, and Pradeep Kumar are the other 15. Ad preference prediction utilizing integrated eeg and sentiment analysis. 2017;92(S1):77-88, Neural Networks.

There were sixteen authors: Goldberg, Nichols, Oki, Oki, and Terry. Weaving a web of knowledge with the help of collaborative filtering. Published in 1992 in Communications of the ACM, Volume 35, Issue 12 (December), Pages 61–70.

Uzay Kaymak, Frederik Hogenboom, Flavius Frasincar, Frank Goossen, and Wouter IJntema make up the group of 17. Using a cf-idf semantic recommender, we can tailor the news to each individual.

On page 10 of the 2011 ACM proceedings for the International Conference on Web Intelligence, Mining, and Semantics.

Number Eighteen: Guibing Guo and Mohamed Elgendi. A novel EEG-based recommender system for 3D e-commerce. 2013;1(1):61-65 in the Journal of Advanced Management Science.

Chia-Cheng Hsu, Hsin-Chin Chen, Kuo-Kuang Huang, and Yueh-Min Huang make up the group of 19 authors. Facebook's artificial bee colony algorithm applied to the problem of recommending supplemental reading material depending on the user's preferred method of learning. Publication date: 2012;64(5):1506-1513 in Computers & Mathematics with Applications.

Twenty-two. Clayton J. Hutto & Eric Gilbert. Vader: Simplifying rule-based models for social media sentiment analysis. To be published in 2014's eblogs and Social Media Conference's

Proceedings. R. K. Katarya, No. 21. Metaheuristic artificial bee movie recommendation system. Research in Neural Computation and Its Applications, 2018, pp. 1-8. We're up to number 22 with Rahul Katarya and Om Prakash Verma. The FCM and Grey Wolf optimizer are used in the recommendation system. Written by Rahul Katarya and Om Prakash Verma, pages 1-9 in 2016's Neural Computing and Applications. An efficient group movie recommendation system using cuckoo search. 2017;18(2):105-112 of the Egyptian Journal of Informatics. Cane WK Leung, Stephen CF Chan, and Fu-lai Chung make up our 24th. A rating inference strategy for combining collaborative filtering with sentiment analysis. To be found in the ECAI recommender systems workshop proceedings, pages 62–66, 2006.

Li Fangfang, Xu Guandong, and Cao Longbing make up the 25th spot. Matrix factorization on two levels for recommendation engines. Hui Li, Jiangtao Cui, Bingqing Shen, and Jianfeng Ma. 2016. Neural Computing and Applications 27(8):2267-2278. Group sentiment analysis in microblogs facilitates a smart movie recommendation system. To cite this article: Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. "Neurocomputing," 210:164-173 (2016). The current state and future directions of content-based recommender systems. Refer to pp. 73–105 of the Recommender Systems Handbook. 2012, Chapter 28, Springer, Steven Loria. The text processing tool Textblob. Walaa Medhat, Ahmed Hassan, and Hoda Korashy, "TextBlob: Simplified Text Processing," 2014. An overview of techniques and software for sentiment analysis. 2014, Vol. 5, Issue 4: 1093–1113, Ain Shams Engineering Journal 30 Prem Melville, Raymond J. Mooney, and Ramadass Nagarajan. Improved suggestions using content-based collaborative filtering. The authors of the 31st entry are Cristiano Nascimento, Alberto HF Laender, Altigran S da Silva, and Marcos Andr'e Gonc alves. Their work was published in Aaai/iaai, 23:187-192, 2002. A method for recommending research papers that doesn't rely on any particular sources.

Between pages 297 and 306 of the Digital Libraries: Eleventh Annual International Joint Conference Proceedings. ACM, 2011. There were 32 of them: Jos'e M. Noguera, Manuel J. Barranco, Rafael J. Segura, and Luis Mart'Nez. A hybrid tourist recommendation system based on mobile 3D GIS. Science of Information, 215(2012), pp. 37-52.

Simon Philip, P. Shola, and A. Ovye make up the 33rd member of the group. Research paper recommendation using a content-based method in a digital library. Advances in Computers and Applications: An International Journal, Volume 5(10), Issue 10, 2014.

Carlos Porcel, Antonio Tejeda-Lorente, Miguel Angel Martnez, and Enrique Herrera-Viedma make up the group of 34. A hybrid recommender system for the targeted distribution of technology transfer office research resources. Published in Information Sciences, 184(1), 2012, pp. 1-19. Rudy Prabowo and Michael Thelwall, who make up our 35th spot. Combining methods for analyzing sentiment. 2009;3(2):143-157 of the Journal of Informetrics.

Ravi, Kumar, and Ravi, Vadlamani, 36. A review of opinion mining, sentiment analysis, and their respective tasks, methodologies, and practical applications. 2015, 89:14–46, Knowledge-Based Systems.

Twenty-seventh: Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens is an accessible framework for group news aggregation and filtering. Computer Supported Collaborative Work Conference Proceedings, pages 175–186. ACM, 1994.

Asad Al Faraby, Bambang Pudjoatmodjo, and Ade Romadhony (#38). A hybrid approach is used in the online shopping recommendation system. Pages 166–169 in Proceedings of the International Conference on Information and Communication Technology. Renata L. Rosa, Demsteneso Z. Rodriguez, and Grac a Bressan. IEEE Conference, 2013. A social network sentiment analysis-driven music recommendation system. 2015, Vol. 61, Issue 3 (pp. 359–367), IEEE Transactions on Consumer Electronics, 40 Alan Said, Alejandro Bellog'n, and Arjen De Vries. An application-based assessment strategy for top-n recommender systems. Presented at the ACM RecSys LSRS Workshop, 2013. Citeseer.

Vivek Sembium, Rajeev Rastogi, Atul Saroop, and Srujana Merugu are at number 41 on the list. Size suggestions for the consumer market. Pages 243–250, Recommender Systems 11th International Conference Proceedings. Yue Shi, Martha Larson, and Alan Hanjalic (No. 42, ACM, 2017). Beyond the user-item matrix, the state of the art and the difficulties of the future of collaborative filtering. Mohammad Soleymani, Guillaume Chanel, Joep JM Kierkels, and Thierry Pun. 2014. ACM Computing Surveys, Vol. 47, No. 1.

Ranking movie sequences based on how they make you feel by analyzing your heart rate and other biometric data.

pages 32–39, 2008, Second Workshop on Multimedia Semantics.

Number 44: Stephan Spiegel. Matrix factorization provides the basis for a hybrid approach to recommender systems. Berlin's Technical University's Department of Agent Technologies and Telecommunications in 2009.

Supiya Ujjin and Peter J. Bentley, both ranked 45. A recommendation engine based on particle swarm optimization.
Pages 124–131 in Swarm Intelligence Symposium Proceedings. IEEE, 2003.
Count them: 46: Robin Van Meteren and Maarten Van Someren. Recommendation based on the analysis of material. Workshop on Machine Learning for the Digital Age, pp 47–56, 2000.
Rapha el Fournier-S'niehotta and Tiphaine Viard, number 47. Content-based and link-stream elements for predicting movie ratings. Reference: 2018; arXiv preprint arXiv:1805.02893.
Toshio Okamoto, 48, and Xin Wan. Supporting group study with a recommender system that continually learns from its mistakes. 20(5):611-621, 2011 issue of Neural Computing and Applications.
MohammedWasid, Vibhor Kant, and the number 49. Fuzzy characteristics are used in a particle swarm algorithm to improve collaborative filtering-based recommender systems. 2015, 54:440-448, Procedia - Computer Science.
Mahendra Yadava, Pradeep Kumar, Rajkumar Saini, Partha Pratim Roy, and Debi Prosad Dogra round up the list's top 50. Neuromarketing based on the study of eeg signals.
2017;76(18):19087-19111 in Multimedia Tools and Applications.
51 people were named, including Chenxing Yang, Baogang Wei, Jiangqin Wu, Yin Zhang, and Liang Zhang.
Cares is a cadaveric recommendation system based on rankings. Referenced on pages 203-212 of the Ninth Joint Conference Proceedings on Digital Libraries. ACM, 2009.
52 - Yuan-Yuan Xu, Xu He, Fan Min, and Heng-Ru Zhang. An adaptive recommender system that takes into account user feedback. 2015 Mathematics of Engineering Problems.
Yao Zhou, Lei Zhang, and Zhang Yi make up our 53rd. Deep neural networks for forecasting a film's earnings at the box office. 2017; page range: 1-11 in Neural Computing and Applications.