

Alleviation of Cold Start in Movie Recommendation Systems using Sentiment Analysis of Multi-Modal Social Networks

R. Ravanmehr¹, M. Mirhasani²

1- Computer Engineering Department, Central Tehran branch, Islamic Azad University, Tehran, IRAN.
(r.ravanmehr@iauctb.ac.ir)

2- Computer Engineering Department, Central Tehran branch, Islamic Azad University, Tehran IRAN.

Received (2020-04-26)

Accepted (2020-12-22)

Abstract: The movie recommendation systems are always faced with the new movie cold start problem. Nowadays, social media platform such as Twitter is considered as a rich source of information in various domains, like movies, motivated us to exploit Twitter's content to tackle the movie cold start problem. In this study, we propose a hybrid movie recommendation method utilizing microblogs, movie features, and sentiment lexicon to reduce the effect of data sparsity. For this purpose, first, the movie features are extracted from the Internet Movie Database (IMDB), and the average IMDB score is calculated during the 7-days opening of the movie. Second, the related tweets of the movie and the cast are retrieved by the Twitter API. Third, the polarity of tweets and the public's feeling towards the target movie is extracted using sentiment lexicon analysis. Finally, the results of the three previous steps are integrated, and the prediction is obtained. Our results are compared with the sales volume of the target movie in 7-days opening, which is available in the Mojo Box office. In addition to the real-world benchmarking, we performed extensive experiments to demonstrate the accuracy and effectiveness of our proposed approach in comparison with the other state-of-the-art methods.

Keywords: Movie Recommendation Systems, Cold Start, Sentiment Analysis, Twitter, IMDB, Mojo Box office.

How to cite this article:

R. Ravanmehr, M. Mirhasani. Alleviation of Cold Start in Movie Recommendation Systems using Sentiment Analysis of Multi-Modal Social Networks. J. ADV COMP ENG TECHNOL, 6(4) Autumn 2020 : 251-264

I. INTRODUCTION

Due to the exponential growth of information in social network services, identifying interesting and useful content from the vast amount of alternatives has become a crucial issue. Recommender systems have been developed to provide suggestions for individual users in various domains from a potentially overwhelming set of choices, including books, movies, songs, websites, or blogs [1]. The cold start is one of the most challenging issues facing any recommender system, which appears due to the lack of sufficient information and ratings

[2]. The daily movies released on online streaming platforms and theaters necessitate the development of solutions to alleviate the cold start of movie recommender systems.

Social media platforms are becoming increasingly mainstream, which provides valuable user-generated information by publishing and sharing content. While social networks and recommender systems were introduced individually, the current state-of-the-art recommendation system relies on the social-media characteristics to personalize recommendations. This infinite environment of users' comments and opinions, making social networks a recognized media for users and of course, researchers. Twitter, Facebook,



This work is licensed under the Creative Commons Attribution 4.0 International Licence.

To view a copy of this licence, visit <https://creativecommons.org/licenses/by/4.0/>

and the Internet Movie Database (IMDB) are examples of these successful and popular social networks. Twitter has become one of the most popular micro-blogging platforms recently. Twitter has evolved to become a source of rich and varied information. This is due to the nature of the microblogs on which people post real-time messages about their opinions on a variety of topics, discuss current issues and express positive/negative sentiments for products they use in daily life. In fact, these reasons motivate us to poll these tweets to get a sense of general sentiment for the new movies which encounter a cold start problem. Twitter has a powerful tool known as a hashtag, which allows to retrieve of tweets based on different topics. Anyone can start using a hashtag or can join a discussion by typing a specific topic and using the "#" sign [3].

Media monitoring is one of the most important concepts used in the domain of social networks and, in particular, Twitter, which is discussed in this study. Media monitoring refers to an activity that monitors social networking channels in collecting user comments and sentiments on a specific topic. Sentiment analysis over Twitter offers organizations a fast and effective way to monitor the publics' positive and negative feelings with the aim of identifying attitudes and opinions that are expressed towards brands, movies, directors, etc. [4]. In this research, we present a new methodology for movie cold start problem from a social media platform, taking into account tweets, IMDB, and movie features. For this purpose, we integrate different social media information to form a multi-modal social network.

Collaborative filtering (CF) is a widely used recommendation technique. This method operates by identifying and specifying similar users or items. For example, if the user likes "The Avengers", then the system suggests the "Iron Man" to him based on other similar users' preferences. Although CF has been used in many large commercial websites, it suffers from ratings' sparsity [5]. Moreover, the cold start of the new item (movie) is always present in this method [6].

Content-based recommender systems alleviate cold start by suggesting items with similar characteristics, but the variation in recommended suggestions sharply drops after a while [7] In several studies, the combined methods have been

used to improve the performance of existing recommender systems [8][9]. In this research, we intend to eliminate the lack of information on a new movie cold start problem using sentiment analysis on the new data source in the form of a hybrid recommender system.

Considering the importance and popularity of social networks, the main innovation of this research is to employ the users' tweets as a rich source of up-to-date user-generated comments, in the absence of enough ratings on a new movie to improve the cold start in movie recommender systems. Contrary to the usual methods in which guessing user votes concerning a target movie is based on previous choices, we will rely on analyzing and predicting the polarity of the extracted tweets about the target movie and its cast. To refine and optimize the use of the related microblogs, the hashtag of the movie title is specified as the search index. For this purpose, using media monitoring of microblogs, we focus on collecting three types of tweets using Twitter API to collect tweets involving user comments on the target movie with a high degree of confidence. Afterward, we not only analyze the sentiments of received tweets but also combine the average IMDB score of the movie with users' sentiment analysis result in an attempt to present the most accurate recommendation to each user. To evaluate the suggested approach, the results will be compared with the information available on Box Office Mojo, which details the sales volume of the movie in the first week of the public release. Box Office Mojo tracks box office revenue in a systematic way and publishes data on its website.

The rest of this paper is organized as follows: In section 2, the previous work to eliminate the cold start problem in recommender systems will be reviewed. Then, the proposed method of this research and its related phases will be presented in section 3. The extensive simulations are carried out to evaluate the performance and accuracy of the proposed approach in section 4. Section 5 includes conclusions and future works.

II. PREVIOUS WORKS

In this chapter, we will consider earlier works on cold start in recommender systems. First, we generally review several methods that have been

provided in the literature to overcome the cold start in recommender systems, and then we focus on the approaches to tackle the problem of cold start in movie recommender systems. For the latter, we divide existing methods into two main categories, as explained in subsections 2 and 3.

1. Cold start in recommender systems

In recent years, the use of recommender systems has been increased with the expansion of e-commerce companies. Amazon and Netflix are examples of companies providing recommendations to the users based on their selected items and previous purchases.

Although several attempts have been made to increase the accuracy of the recommender systems, the problem of cold start still exists. In some papers, researchers have made efforts to compensate for this lack of information by asking a series of questions concerning user interests or suggesting several items [10][11]. There have been studies to use statistical information rather than a past vote of a user, and it is assumed that users in the same geographic region will probably have the same tastes and will assign identical rates to items [12]. Pandey et al. suggested to calculate the similarity between users by considering the statistical data of new and old users [13]. In some other studies such as [14], items are firstly clustered according to the user-item matrix, and then the content information of the item and clustering results are combined to obtain the decision tree, which results in the association between the new and the current items. Tobías et al. evaluated a number of matrix factorization models for cross-domain collaborative filtering that leverage metadata as a bridge between items liked by users in different domains [15].

A hybrid recommender system has been proposed by Katarya et al., which utilized k-means clustering algorithm with a bio-inspired artificial bee colony (ABC) optimization technique [16]. Guo et al. proposed a TrustSVD technique that is an extension of SVD++ with social trust information based on a novel trust-based matrix factorization model that incorporated both rating and trust information [17]. This article takes into account both the explicit and implicit influence of ratings and trust information when predicting ratings of unknown items.

Zhang et al. utilized a probabilistic neural

network as a learning method, which calculates the level of trust among users, and a sparse matrix modifies the votes by predicting unregistered votes [18]. A confidence-based Linear Frequency Modulated (LFM) method such as SVD++ has been applied to predict user votes on a new item in [19]. Using ontological concepts that have been widely considered in computer science and various fields of data storage and retrieval in recent years, several papers suggest a trend-based approach by constructing an ontology model to collect all available information from items assigned to the new user as an active item[20].

Recently, a systematic literature review on the CF-based recommender system that employs social network data to mitigate the cold-start problem has been published by Camacho et al. [2].

2. Cold start in movie recommender systems using social network data

Although social networks and recommender systems are independently introduced and used, recently, successful recommender systems are employing the contents of social networks to generate personalized suggestions for users.

Reshma et al. proposed a new approach for the prediction of votes for items which are based on the directed and transitive with timestamps and profile similarity from the social network along with the user-rated information [21]. This approach utilizes the trust concept in a very limited way and does not consider trust beyond the second level. So, it only covers a limited audience.

Revathy et al. developed an architecture combining social networks like Facebook, Twitter, Pinterest, and contextual data to overcome the user cold start problem and handled the new item cold-start problem using Wikipedia, or some other public domains [22]. The Collaborative Filtering method is enhanced through clustering using contextual information. Facebook social network helps to face the cold start and produces recommendations by using Facebook pages of movies and fans. The combination of the similarity results obtained from Movie Rating Systems and the movie "Facebook Page" information have been used to solve the cold start problem [23]. Lee et al. have only extracted "Like" and "Co-Like" on "Facebook Page", and have not employed the

pool of Facebook information, which contains users' activities such as post, comment, or respond. Similar to [23], Rosli et al. have utilized the Facebook fan pages information such as "Likes" or "Dislikes" to compute the scores of the recommended products [24]. The Singular-Value Decomposition (SVD) was applied for dimension reduction to enable more-efficient computations.

Ji et al. proposed a hybrid method that uses user-item ratings to build a content association between users and items based on three factors-user interest in selected tags, tag-keyword relation, and item correlation with extracted keywords, and then recommends the items with high similarity in content to users.

IMDB, also known as the Internet Movie Database, is the most reliable source for movies, TV series, and celebrities' information. The IMDB users' comments are reflected through writing on movies, presenting numerical statistics, or selecting movie stars. Illig et al. attempted to estimate the popularity of the movie by tracking the tweets containing the title of a movie, searching users' comments about the target movie on IMDb, and forming the sentiment graph linking named-entities [7].

The Twitter social network has also been used in some papers to cope with cold start. Natarajan et al. utilized the Twitter profile of a new user to collect information about him/her. In this regard, the users' interests and preferences are extracted by analyzing their tweets [26].

Moshfeghi et al. applied sentiment analysis on users' ratings and reviews to calculate their degree of satisfaction to better deal with cold start of a new item or user [27]. The proposed framework relied on an extension of Latent Dirichlet Allocation (LDA), and on gradient boosted trees for the final prediction. The required data have been extracted from IMDB (the movie plot summary and the reviews) to define the semantic and emotional spaces in 3 different categories: actor, director, and genre. However, utilizing emotions extracted from other popular social media sources such as Twitter or Facebook is an important gap that has not been addressed.

The problem of recommending new movies has been addressed in [12] by monitoring information from social media services. More specifically, it focused on exploring how the reputation of movies, directors, and actors can be used to tackle

this issue. For this purpose, the authors obtained users' comments from Twitter for directors and actors of 60 new movies, finalists on five popular movie awards ceremonies. In fact, the proposed method in [12] collects tweets about the movie's director and cast that do not belong specifically to the considered movie, and the performance of actors in other films is also considered. For example, an actor may have received positive comments in many movies, but has not succeeded in the specified movie. As a result, the analysis of users' opinions for the movie's cast is closely related to the cast's performances in the previously released movies, which makes it difficult to obtain a precise result from the sentiment analysis of the users' comments for the movie's cast in the new-released film.

3. Cold start in movie recommender systems using movie features

In several types of research, a series of movie features have been employed as the main sources to tackle the cold start-up problem. These methods do not mainly use information from social networks.

The main goal of some approaches is to improve the collaborative filtering in recommender systems to generate more accurate recommendations. For example, it is attempted to find items that are most similar to the target item [28]. Hence, using the cosine similarity, the similarity of movies, and the selection of appropriate neighbors for accurate prediction of votes are discussed. Deldjoo et al. introduced a new movie recommender system that addresses the new item cold start problem in the movie domain by integrating state-of-the-art audio and visual descriptors [29]. These descriptors can be automatically extracted from video content and constitute what they call the movie genome and exploiting an effective data fusion method named canonical correlation analysis.

Yi et al. concentrated on solving the cold start, creating a bridge between common features of movies (such as director and actors) and finding a set of similar movies [30]. The main aim of this approach is to bridge the gap between movie labels and movie similarities. The authors attempt to optimize the movie similarity measure by computing the similarities among directors and actors. They claimed that the new movie in the

cold start scenario has been improved by the use of movie labels for new movies to find a similar set.

Recently, several attempts have been made to tackle the cold start problem in collaborative filtering. One such attempt used category correlations of contents. Choi et al. proposed a recommendation system employing the movie genre correlations [11]. Since the movie genres are described by experts such as directors or producers, they can be more reliable than user ratings. Thus, the approaches that employ genre correlations for movie suggestions can be more accurate. The authors analyzed genre correlations according to the number of genres and also to the decade when the movie was made. A novel way of integrating movie content embeddings in CF has been proposed by Anwaar et al. [6]. The developed framework (HRS-CE) generates the user profiles that depict the type of movie content in which a particular user is interested. The higher representation for a movie description, obtained using content embeddings, is combined with similarity techniques to perform rating predictions.

Pirasteh et al. also attempted to recommend movies with maximum possible similarity to what the user has watched in the past [31]. This technique involves user tastes in this interaction and identifies the features of movies that have been repeatedly watched by the user.

Table 1 shows a summary of the methods discussed in sections 2 and 3, which are directly about the cold start problem in the movie recommender system. Our goal in this study is to employ the features of a movie to identify the user's preferences. However, the discovery of the user's taste is not limited by the number of times the users watch certain movies (with features such as the same director or actor) [31]. Indeed, this recognition is based on analyzing the tweets posted by users on movies with the same features. In fact, our work is devoted to identifying movie features in the context of comments published about the movie on Twitter. Similar to [31] we take advantage of movie features such as actors and directors, but we will look for comments on these features in user tweets and, like [20] we use sentiment analysis techniques in an attempt to analyze users' feedback through short posts shared by them in the form of microblogs on

Twitter.

Moreover, by referring to IMDB, we extract the available information about movie features, and average user votes on the target movie in the first seven days of its release using IMDB API. The results from the above steps are calculated with different weights to find the target movie score. Now, we can predict to what degree the movie will be appealed to the users interested in that particular genre. Unlike almost all methods discussed in this section, our prediction results are compared with the sales volume of the target movie in 7-days opening, which can be retrieved through the third data source in our proposed system, Mojo Box office.

TABLE 1
COMPARISON OF PREVIOUS WORKS IN COLD START PROBLEM OF MOVIE RECOMMENDER SYSTEM

Paper	Method	Social Media Source	Movie Feature Source	QoS Metrics
[22]	Integrating social networks and contextual data to overcome cold start	Facebook, Twitter	-	-
[12]	Movies popularity is deduced from social-media trends related to the corresponding new movie	Twitter	-	MAE F-M
[21]	Combining the CF with the social behavior of users for alleviating sparsity and cold start problems	General SNS	-	MAE Coverage
[26]	User's interests and preferences are extracted from their Twitter profile using sentiment analysis	Twitter	-	Accuracy
[27]	Sentiment analysis applied to user ratings and reviews to calculate their degree of satisfaction using LDA extension	IMDB	-	MAE MSE
[24]	Using Facebook Fan Pages information such as "Likes" or "Dislikes" to compute the scores of recommended products	Facebook	-	MAE Coverage
[23]	Integrating the Facebook Fan Page data and the genre-classifications data from Yahoo!	Facebook	-	Accuracy
[25]	Incorporating content-based information and social information, recommendation is performed for the items with high similarity in content	Weibo	-	MAE Recall
[29]	Integrating audio and visual descriptors, exploiting an effective data fusion method, and proposing a two-step hybrid approach on movie genome to recommend cold items.	-	Movie genome (genre, cast, ...)	Relevance Novelty Diversity
[30]	Creating a bridge between common features of movies (such as director and actors) and finding a set of similar movies	-	Genre, Movie Director	Precision Recall
[11]	Utilizing genre correlations to cover the RS shortcomings	-	Genre	Standard deviation
[28]	Finding items that are most similar to the target item in order to improve CF	-	Movie Title	MAE
[31]	Measuring the similarity between items by utilizing the genre and director of movies	-	Genre, Movie Director	MAE RMSE
[6]	Integrating movie content embedding in collaborative filtering	-	Genre, Movie Director	MAE RMSE

III. THE PROPOSED METHOD

In this section, a hybrid approach is proposed to alleviate the cold start in movie recommender systems using sentiment analysis of extracted blogposts. Generally, our approach consists of four phases. In the first phase, social-media monitoring is performed and the required information is collected from the related social networks (Twitter and IMDB in this study). Then, in the pre-processing phase, the collected tweets are cleaned, classified, and stored in JSON format. In the third step, after extracting tweets features in 3 different categories, sentiment analysis is performed based on a Lexical and Emoji resource dictionary to identify the polarity of tweets related to target movie, and finally, the target movie score is calculated based on previous phases' information.

We have employed different types of APIs to collect the required information from different social media sources. For this purpose, we use Twitter API to collect 3 types of tweets in the first step of the proposed approach. The main index term for collecting related tweets is the hashtag of movie title as well as the name of the main cast and film director. By employing IMDB APIs, we receive the features and information about the new movie that suffers from cold start problem. Our required information includes the names of the actors, the director, the genre, and the rating of users of IMDB to the desired movie.

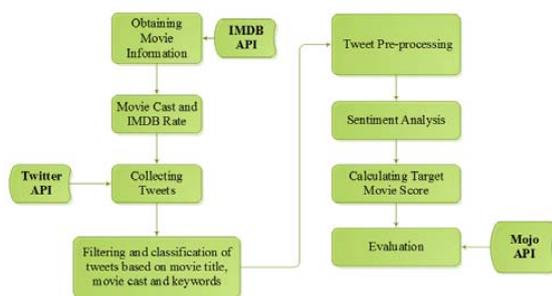


Fig. 1. Proposed approach

Moreover, we utilize the Box Office Mojo API to obtain the movie sale volume on the opening weekend after public release. For the sentiment analysis purpose, we utilize two different dictionaries. The first one is SentiWordNet 3.0 [32] an enhanced lexical resource explicitly devised for

supporting sentiment classification and opinion mining. The second one is an online Emoji dictionary from "emoji-emotions" [33]. Based on these dictionaries, the positive and negative poles of users' tweets are obtained to specify the degree of satisfaction or dissatisfaction.

The outcome of the fourth step is obtained by putting together the parameters that consist of the result of tweets' sentiment analysis, the average IMDB score, and the popularity estimation of the movie cast. This result predicts the likelihood of satisfaction of a cold start movie.

1. Collecting Tweets

The first issue is to employ suitable tools to have access to microblogs. Twitter's API is among the tools that are extensively applied in this research to find the tweets. In addition, there are two other APIs, namely mojo API and IMDB API. The former is used to obtain movie box office information, and the position of the movie in the weekly chart, and the latter is responsible for collecting the movie features, attributes, and IMDB rating.

First, the movie features are extracted from the IMDB, and the average IMDB score is calculated during the first seven days of the movie release. Then, the tweets based on movie features and special keywords are retrieved and extracted into three types, which are introduced below.

However, before we start the discussions about the classification of collected tweets, it should be noted that users usually utilize a few specific words to express their interest or dissatisfaction with movies on social networks; for example, if we consider the following tweets collected using Twitter API:

```

[created_at] => Sun Jul 09 18:52:04 +0000 2017
[id] => 8.841230026439E+17
[id_str] => 884123002643902465
[text] => @MUTGuru Just watched #SpiderManHomecoming. It was fantastic. Best spiderman movie coming from a fan since childhood.

[created_at] => Sun Jul 09 07:09:18 +0000 2017
[id] => 8.8394614422871E+17
[id_str] => 883946144228712452
[text] => I've seen #SpiderManHomecoming already It was boring ☹ @MarvelStudios
  
```

Fig. 2. Example of tweets received using Twitter API

It can be seen that users usually express their opinions about a movie in the text after the phrases "just watched" or "I've seen" (Figure 2). For this purpose, we have considered a list of these special phrases that are used to collect and analyze the related tweets. Now, taking into account the above considerations, the tweets are retrieved and extracted in three different types. The first category is the tweets that have been collected using the hashtag of movie title plus "keywords" (T1). For example, when the "watched" or "seen" is searched with the hashtag of the movie title, we obtain the tweets posted after the user has watched the movie, which expresses their opinion about the movie with a high degree of confidence.

The data extracted from IMDB are used together with the hashtag of the movie title to collect the second type of tweets (T2). This information includes the name of the cast and the director of the movie, which is available using IMDB API.

The third category is the tweets that are only collected using the hashtag of the movie title (T3). Considering that we have already focused on the targeted searches for tweets, we collect a part of tweets from this type to prevent the occurrence of locality and any possible bias. The different types of tweets are depicted in Figure 3.

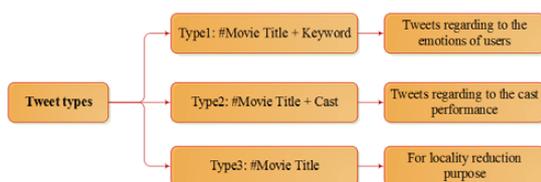


Fig. 3. Classification of collected tweets

```

Input Tweets← The list of tweets retrieve from Twitter API
1. For all  $T_i, T_j \in$  Tweets Do;
2.  $T_i\_Date \leftarrow$  Get date of tweet,  $T_i\_ID \leftarrow$  Get ID of tweet;
3.  $T_j\_Date \leftarrow$  Get date of tweet,  $T_j\_ID \leftarrow$  Get ID of tweet;
4. If  $T_i\_Date == T_j\_Date$ 
   {
     Compare ( $T_i\_ID, T_j\_ID$ );
     If ( $T_i\_ID == T_j\_ID$ )
       Delete  $T_j$ 
       Add Tweets←  $T_i$ 
     Else
       Add Tweets←  $T_i, T_j$ 
   }
Else
  Add Tweets←  $T_i, T_j$ ;
  
```

Fig. 4. Pseudo-code for removing duplicated tweets in the pre-processing phase

2. Pre-processing

In this phase, simple processing is performed on the collected tweets to remove the duplicated ones (Figure 4). The set of tweets receiving from Twitter API are stored in JSON format, and in addition to user tweets, contains information such as time, publishing date, the ID of tweet author, and several other entries. Using the date and ID fields, the duplicated tweets by the same user are eliminated. Therefore, in the pre-processing phase of our approach, the tweets are collected, classified (according to the types explained in section 3.1.), and sorted one-by-one to construct the dataset with respect to each movie.

3. Sentiment analysis

Sentiment analysis over Twitter provides a fast and effective way to monitor the public's feelings toward specific subjects such as movies. We conclude that the movie is more successful when the number of positive tweets is larger than the negative ones. For example, tweets that include the **great** and **amazing** attributes receive the value of 1 and have a positive polarity, but the tweets involving negative attributes such as **bad**, **terrible**, **awful**, and so forth receive the value of -1. These attributes are chosen from SentiWordNet 3.0 dictionary [32], and since it does not include slangs and Emoji, online dictionaries are also employed [33]. Therefore, the number of analyzed tweets and their values are counted, and the degree of satisfaction and dissatisfaction of users whose tweets are collected are estimated using in "Eq. (1)":

$$\text{Sent}_{\text{rate}} = \frac{N_p - N_n}{N_t} \quad (1)$$

N_p , N_n , and N_t represent the number of positive, negative, and total tweets.

As explained in section 3.1. T_1 includes tweets based on the hashtag of movie title plus the keywords. For this purpose, we consider 13 keywords that directly express the user's opinion with certain polarity about the movie, such as love, great, terrible, etc. Table2. Therefore, the number of extracted tweets with the mentioned hashtag (#MovieTitle+Keyword) that directly reflect the user satisfaction (or dissatisfaction) is substituted in "Eq. (1)". Moreover, we have considered 10 other keywords, such as watched, seen, etc. to collect the tweets of user comments about the target movie without certain polarity. The use of these keywords, along with the movie title, makes it possible to review and analyze the tweets that reflect the users' comments. The sentiment analysis is performed on these tweets, and after determining their polarity, their values, along with the results of direct tweets, will be substituted in "Eq. (1)".

TABLE 2
HASHTAGS TO COLLECT T1 TYPE TWEETS

Num.	Hashtags for collecting tweets with certain polarity: #MovieTitle+Keyword	Hashtags for collecting tweets require sentiment analysis: #MovieTitle+Keyword
1	#MovieTitle + love	#MovieTitle + watch
2	#MovieTitle + great	#MovieTitle + watched
3	#MovieTitle + wow	#MovieTitle + watching
4	#MovieTitle + omg	#MovieTitle + see
5	#MovieTitle + ☺	#MovieTitle + saw
6	#MovieTitle + good	#MovieTitle + seen
7	#MovieTitle + amazing	#MovieTitle + seeing
8	#MovieTitle + bad	#MovieTitle + go
9	#MovieTitle + terrible	#MovieTitle + gone
10	#MovieTitle + boring	#MovieTitle + went
11	#MovieTitle + ☹	
12	#MovieTitle + awful	
13	#MovieTitle + worst	

Information in Table 3 shows the user tweets about the target movie cast (called T2 Type). The names of the cast are obtained from IMDB API. Since an actor may have a brilliant performance in a movie but does not appear so in several other films, the use of tweets assessing the performance of actors and director in a particular film can

increase the accuracy of predictions.

TABLE 3
HASHTAGS TO COLLECT T2 TYPE TWEETS

Hashtags for collecting tweets require sentiment analysis:
#MovieTitle +Cast
#MovieTitle +Cast1
#MovieTitle + Cast2
....
....
#MovieTitle + Director

Using IMDB API, the alternative names (that a movie might have) are obtained in Table 4. As explained in section 3.1, to avoid the occurrence of locality, we will collect a group of tweets (called T3 type) based on Table 4.

TABLE 4
HASHTAGS TO COLLECT T3 TYPE TWEETS

Hashtags for collecting tweets require sentiment analysis:
#MovieTitle
#MovieTitle
#MovieTitle2
#MovieTitle + film
#MovieTitle + movie

4. Target Movie Score

In our approach, the results of user satisfaction of the target movie based on the performed sentiment analysis of T_1 and T_3 tweet types are called Movie Tweet Rate (MTR) "Eq. (3)". The value obtained from the sentiment analysis of T_2 tweet type is considered as the popularity of the movie cast, namely Cast Tweet Rate (CTR) "Eq. (4)".

Now, we calculate the target movie score in the below "Eq. (2)" by combining MTR, CTR, and average IMDB score of the movie during the first seven days of movie release:

$$W_f = \alpha(\text{MTR}) + \beta(\text{CTR}) + \lambda(\text{IMDB Score}) \quad (2)$$

We have considered three weighting factors as α , β , and λ . Since the sentiment analysis of tweets about the movie is the most important component in our approach, the value of α will be higher than the other two values, followed by β and finally λ . The most appropriate values are determined through different analysis have been done and reported in Figure 6 and Table 10 in

section 4.4. According to “Eq. (1)”

$$MTR = \frac{N(\text{positive movie tweets}) - N(\text{negative movie tweets})}{N(\text{total movie tweets})} \tag{3}$$

$$CTR = \frac{N(\text{positive cast tweet}) - N(\text{negative cast tweet})}{N(\text{total cast tweet})} \tag{4}$$

After obtaining the movie score based on W_f , we consider movie sale volume for 7-days opening box office, as the accuracy indicator for the results of tweets sentiment analysis. Now, if the value of W_f is high, it is expected to have considerable movie box office revenue during the weekend, and the movie should appeal to the audience. Therefore, by utilizing the target movie score, an effective recommendation can be placed for the newly released movie that is faced with the cold start problem.

Now, let's take the "SpiderMan: Homecoming" movie as an example. The movie was released on July 7, 2017, in cinemas. The tweets related to this movie were received from this date up to the next seven days. After collecting tweets, it is time to analyze the tweets and obtain the polarity of them. We have received 25,200 tweets per day for the SpiderMan movie.

TABLE 5
SENTIMENT ANALYSIS OF COLLECTED "SpiderMan: Homecoming" CAST TWEETS ON THE FIRST DAY OF RELEASE FOR T1 TYPE

		T ₁ tweet type										
Polarity		#SpiderMan: Homecoming +watch	#SpiderMan: Homecoming + watching	#SpiderMan: Homecoming +watched	#SpiderMan: Homecoming +sec	#SpiderMan: Homecoming +saw	#SpiderMan: Homecoming +seeing	#SpiderMan: Homecoming +seen	#SpiderMan: Homecoming +go	#SpiderMan: Homecoming +went	#SpiderMan: Homecoming +going	sum
	Positive	192	408	93	283	411	336	497	184	165	139	2708
Negative	19	14	30	24	20	28	14	62	8	13	232	

This trend continued during the first seven days of releasing the movie, and at the end of this period, the number of tweets has been reached to

176,400. The collected tweets were examined at the pre-processing step. Then, sentiment analysis is performed on T1 and T3 types of tweets, as shown in Table 5 and Table 6, respectively.

TABLE 6
SENTIMENT ANALYSIS OF COLLECTED "SpiderMan: Homecoming" CAST TWEETS ON THE FIRST DAY OF RELEASE FOR T3 TYPE

		T ₃ tweet type				
Polarity		#SpiderMan: Homecoming	#SpiderMan	#SpiderMan: Homecoming +Movie	#SpiderMan: Homecoming +film	Sum
	Positive	249	153	379	220	1001
Negative	15	20	6	42	83	

In addition to T1 and T3, the sentiment analysis is performed on the T2 tweet type, which is about the "SpiderMan: Homecoming" movie cast. For this purpose, first, we retrieve target movie cast information using IMDB API. Then, after collecting the related tweets for movie cast, we employ SentiWordNet 3.0 for sentiment analysis to obtain the polarity of users' tweets. The results of this analysis have been shown in Table7.

TABLE 7
SENTIMENT ANALYSIS OF COLLECTED "SpiderMan: Homecoming" CAST TWEETS ON THE FIRST DAY OF RELEASE FOR T2 TYPE

		T ₂ tweet type						
Polarity		#SpiderMan: Homecoming+ RobertDowneyJr	#SpiderMan: Homecoming+ TomHolland	#SpiderMan: Homecoming+ MichaelKeaton	#SpiderMan: Homecoming+ MarisaTomei	#SpiderMan: Homecoming+ Zendaya	#SpiderMan: Homecoming+ JonWatts	Sum
	Positive	577	263	229	369	446	283	2167
Negative	33	38	25	43	42	46	227	

Now, using IMDB API, we retrieve the movie rating base on the users' votes on the IMDB

website and calculate the average movie rating within the first week of the movie release (Table 8).

TABLE 8
MOVIE RATING FOR "SpiderMan: Homecoming" IN IMDB

Date	Vote	IMDB rate
2017.07.07	25,625	8.3
2017.07.08	38,053	8.2
2017.07.09	49,745	8.2
2017.07.10	50,386	8.2
2017.07.11	63,484	8.2
2017.07.12	76,890	8.2
2017.07.13	88,315	8.1

IV. EVALUATIONS

In this section, we describe our experimental evaluation and discuss the results of the proposed approach.

1. Datasets

As explained in the previous chapter, three types of datasets have been used in this research: the movie features extracted from IMDB, the weekly box office results from Box Office Mojo, and the collected tweets related to the target movies using Twitter API. To ensure that the tweets collected are related to the target movie as much as possible, the movie title has been used in the form of the hashtag "#" as the main search index. The research is conducted on eighteen films released in 2016-2017. In total, over 3,000,000 tweets have been collected in three different categories (T1, T2, and T3) for sentiment analysis purposes. Using sentiment word dictionary, it has been attempted to analyze the satisfaction or dissatisfaction of viewers after watching the movie. It should be noted that several well-known movie datasets such as MovieLens could not be employed in this research since in our approach we have attempted to solve the cold start problem for the newly released movie that users' rating has not been registered in these types of datasets. In other words, our approach only relies on the data that has been extracted from real-world media sources such as Twitter, IMDB, and Box Office Mojo.

2. Evaluation metrics

We have considered two different types of metrics for the assessment of the proposed

approach. First, the precision "Eq. (5)" and recall "Eq. (6)" metrics are used to assess the collection and analysis of the tweets:

$$Precision = \frac{true\ positive}{all\ positive\ result(true\ positive + false\ positive)} \quad (5)$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (6)$$

The precision criterion estimates the number of movie-related collected tweets containing user comments, and recall determines how many tweets collected in the form of user comments actually reflect their opinions about the movie.

Second, to investigate the accuracy of the suggested approach to alleviate the problem of cold start of a new movie, we use the Mean Absolute Error (MAE) "Eq. (7)" and Root Mean Square Error (RMSE) "Eq. (8)" metrics:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - p_i| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - p_i)^2} \quad (8)$$

X_i represents the position of the i^{th} movie in the descending order list based on movie score value (W_i). For this purpose, W_i is obtained for all movies studied in this research and is arranged in the descending order. P_i is also the rank of the i^{th} movie based on the order of the box office revenue on the chart among all the movies reviewed in this research. To determine the accuracy of the proposed method, the distance between X_i and P_i is calculated and evaluated.

3. Tweets collection analysis

As explained in section 3.3, three different types of tweets are collected about the target movie. Now, we want to show the effect of these tweets on the precision and accuracy parameters. We first review the effect of T_3 tweets on precision and recall because it shows the popularity of the

movie (first row in Table 9). Indeed, when the release of a movie makes a lot of tweets, it usually indicates the popularity of the movie. Although, it should be noted that a large number of these tweets may contain promotional videos, fan art, soundtrack of a movie, and so on. For example, "Her" is the title of a movie in 2013, which is also used as a personal pronoun in numerous tweets, or "Assassin's Creed", which is a set of awesome multiplayer games. Therefore, preprocessing is required to eliminate irrelevant duplicate tweets in T_3 type. The second type of tweets (T_2) is also considered to improve accuracy (second row in Table 9). The final step is to add type 1 tweets (T_1) to focus on collecting tweets based on "keywords", which will lead to the tweets through which the users are likely to express their opinions on the target movie (third row in Table 9).

As a result, we can collect those tweets ($T_1+T_2+T_3$) with maximum relevance to target movie from among a large number of tweets to perform sentiment analysis on more accurate tweets.

TABLE 9
PRECISION AND RECALL ASSESSMENT FOR DIFFERENT TYPES OF TWEETS

Tweet Type	Precision	Recall
T_3	%51	%58
T_2+T_3 + preprocessing	%66	%74
$T_1+T_2+T_3$ + preprocessing	%79	%91

4. Movie score assessment

As we have shown in "Eq. (2)" the final score of the target movie (W_f) results from the combination of three parameters of MTR, CTR, and average IMDB score, and the corresponding coefficients α , β , and λ . To determine the most effective parameters on W_f between MTR, CTR, and IMDB score, the MAE and RMSE have been calculated in Figure 5.

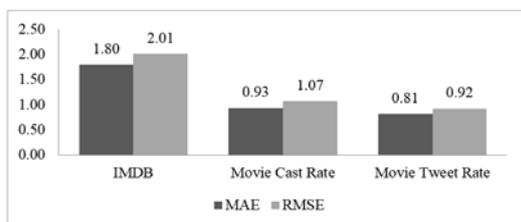


Fig. 5. Comparison of the parameters used in W_f

As shown in Figure 5, the lowest MAE and RMSE values belong to the MTR parameter. Indeed, it shows that the shared tweets about the movie (T_1 and T_3 tweet types) have produced the most accurate results compared to the box office performance of the movie. After the MTR parameter, the most effective parameters on W_f are CTR and IMDB score.

Although the IMDB score parameter increases the error in some special cases, this parameter adjusts the accuracy of the results of W_f and removes the biases about the movie.

For example, releasing the new movies can create an initial excitement, and result in emotional tweets. In these circumstances, the parameters such as IMDB score produce a more realistic result. Now, we calculate the best weight coefficients α , β , and λ in "Eq. (2)".

According to the above discussions, we consider $\alpha > \beta > \lambda$ to ensure that the MTR and then the CTR parameter imposes the most effect on W_f . For this purpose, we calculate W_f for different sets of α , β , and λ for all movies in our dataset to find the optimum set. For this purpose, we compare the results with the box office performance of the movie and calculate MAE in Figure 6.

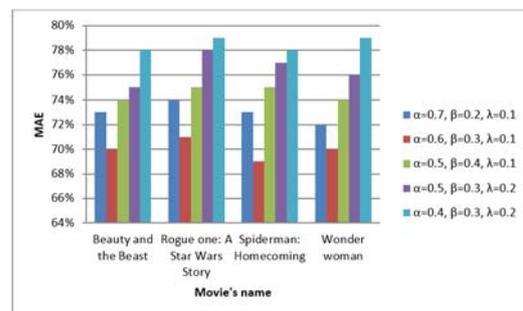


Fig. 6. Comparison of the different weight used in W_f

As shown in Table 10, the average value of movie scores is optimum when the weights are equal to $\alpha=0.6$, $\beta=0.3$, and $\lambda=0.1$ (W_{f2}). It should be mentioned that we have performed the same process for all movies in our dataset to make sure about the values of coefficients α , β , and λ .

TABLE 10
 MOVIE SCORE (WF) ASSESSMENT FOR DIFFERENT WEIGHTS OF α , β AND λ

Movie	W_f	W_{f1}	W_{f2}	W_{f3}	W_{f4}	W_{f5}
		$\alpha=0.7$ $\beta=0.2$ $\lambda=0.1$	$\alpha=0.6$ $\beta=0.3$ $\lambda=0.1$	$\alpha=0.5$ $\beta=0.4$ $\lambda=0.1$	$\alpha=0.5$ $\beta=0.3$ $\lambda=0.2$	$\alpha=0.4$ $\beta=0.3$ $\lambda=0.2$
Beauty and the Beast		0.88	0.87	0.86	0.85	0.85
Rogue one: A Star Wars Story		0.82	0.81	0.81	0.81	0.79
Spiderman: Homecoming		0.84	0.84	0.83	0.83	0.83
Wonder woman		0.87	0.87	0.87	0.86	0.85

5. Comparisons

For more investigations on the results obtained from the proposed approach, it is compared with different methods for solving the cold start problem such as GIS-GD [31] LFM [33] and TrustSVD [17]. All these methods focus on solving the cold start of a new movie in the recommender system and have been discussed in detail in Section 2.

As shown in Figure 7, GIS-GD, LFM, TrustSVD, and our approach have the MAE values of 0.80, 0.84, 0.80, and 0.68, respectively.

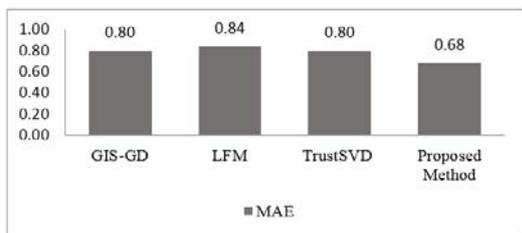


Fig. 7. Comparison of the proposed approach with other methods based on MAE

Moreover, as shown in Figure 8, GIS-GD, TrustSVD, LFM, and our approach have the RMSE values of 0.86, 1.04, 1.04, and 0.84, respectively.

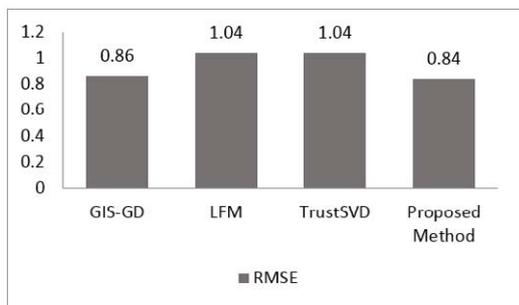


Fig. 8. Comparison of the proposed approach with other methods based on RMSE

In [31] the movie features have been employed to overcome the cold start of the new movie. For this purpose, the movie recommendation system focuses on the similarities between the genre and the director of the target movie and the previously available information. Therefore, the suggested movies have less novelty and serendipity, and on the other hand, the performance of the director and the cast of the target movie are not considered. While in our approach, the similarities between the target movie and the previous movies, and moreover, the performance of the cast in the previous movies, has no effect on the recommendations for the target movie. It should be mentioned that an actor may have shown brilliant performance in previously released movies but does not appear so in the newly released movie that is faced with the cold start problem.

Trust-based methods such as [19] and [17] require prior users' votes, and then matrix sparsity is improved using different techniques such as SVD++ and neighborhood model [19] or weighted-regularization [17]. While in our proposed approach, the cold start problem of the new movie is alleviated employing the sentiment analysis of users' opinions which are continually updated by a wide range of users in social networks such as Twitter.

V. CONCLUSION AND FUTURE WORKS

In this paper, we proposed our model to alleviate the movie cold start problem taking into account users' tweets and movie features. For this purpose, three types of tweets were collected employing the Twitter API using the hashtag of the movie titles as well as a series of keywords for better assessment of the related tweets, and information on the movie features such as the cast, director, and the movie rating on IMDB. The polarity of all the tweets was calculated to highlight the popularity of the movie and its cast. The final target movie rating has been concluded by combining movie tweet rate, cast tweet rate, and average IMDB score of the movie during the first seven days of the movie release. Then, we considered movie sale volume for 7-day opening box office, as the accuracy indicator for the results

of the final target movie rating. For this purpose, we utilized the information available on Box Office Mojo, which details the sales volume of the movie in the first week of the public release. The results of experiments for different movies demonstrated the accuracy and effectiveness of the proposed method in comparison with the other state-of-the-art methods dealing with the cold start of a new movie.

Our approach can be easily extended to identify the polarity of complex sentences in any microblog platforms except Twitter to provide more powerful analysis tools. Since this study's focus was on improving the cold start of a new movie, this fundamental idea can be entirely used for solving the cold start problem in other areas than movies.

As future research, it is suggested to use more powerful lexicon-based and machine learning techniques to increase the accuracy of the sentiment analysis results for polarity identification of complex sentences. Since every movie has several genres, involving more genres with different weights can help to increase the accuracy of recommender system suggestions. Considering the recent advancement in deep neural networks, this approach can be enhanced by employing these networks in feature extraction or rating prediction.

REFERENCES

1. Son, L.H., 2015. HU-FCF++. *Engineering Applications of Artificial Intelligence*, 41(C), pp.207-222.
2. Camacho, L.A.G. and Alves-Souza, S.N., 2018. Social network data to alleviate cold-start in recommender system: A systematic review. *Information Processing & Management*, 54(4), pp.529-544.
3. Otsuka, E., Wallace, S.A. and Chiu, D., 2016. A hashtag recommendation system for twitter data streams. *Computational social networks*, 3(1), p.3.
4. Khan, F.H., Bashir, S. and Qamar, U., 2014. TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision support systems*, 57, pp.245-257.
5. Yang, X., Guo, Y., Liu, Y. and Steck, H., 2014. A survey of collaborative filtering based social recommender systems. *Computer communications*, 41, pp.1-10.
6. Anwaar, F., Iltaf, N., Afzal, H. and Nawaz, R., 2018. HRS-CE: A hybrid framework to integrate content embeddings in recommender systems for cold start items. *Journal of computational science*, 29, pp.9-18.
7. Illig, J., Hotho, A., Jäschke, R. and Stumme, G., 2007. A comparison of content-based tag recommendations in folksonomy systems. In *Knowledge Processing and Data Analysis* (pp. 136-149). Springer, Berlin, Heidelberg.
8. Alahmadi, D.H. and Zeng, X.J., 2015, November. Twitter-based recommender system to address cold-start: A genetic algorithm based trust modelling and probabilistic sentiment analysis. In *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 1045-1052). IEEE.
9. Dang, T.T., Duong, T.H. and Nguyen, H.S., 2014, December. A hybrid framework for enhancing correlation to solve cold-start problem in recommender systems. In the *2014 Seventh IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)* (pp. 1-5). IEEE.
10. Aharon, M., Anava, O., Avigdor-Elgrabli, N., Drachler-Cohen, D., Golan, S. and Somekh, O., 2015, September. Excuseme: Asking users to help in item cold-start recommendations. In *Proceedings of the 9th ACM Conference on Recommender Systems* (pp. 83-90).
11. Choi, S.M., Ko, S.K. and Han, Y.S., 2012. A movie recommendation algorithm based on genre correlations. *Expert Systems with Applications*, 39(9), pp.8079-8085.
12. Santos, J., Peleja, F., Martins, F. and Magalhães, J., 2017, October. Improving cold-start

recommendations with social-media trends and reputations. In International Symposium on Intelligent Data Analysis (pp. 297-309). Springer, Cham.

13. Pandey, A.K. and Rajpoot, D.S., 2016, December. Resolving cold start problem in recommendation system using demographic approach. In 2016 International Conference on Signal Processing and Communication (ICSC) (pp. 213-218). IEEE.

14. Sun, D., Luo, Z. and Zhang, F., 2011, October. A novel approach for collaborative filtering to alleviate the new item cold-start problem. In 2011 11th International Symposium on Communications & Information Technologies (ISCIT) (pp. 402-406). IEEE.

15. Fernández-Tobías, I., Cantador, I., Tomeo, P., Anelli, V.W. and Di Noia, T., 2019. Addressing the user cold start with cross-domain collaborative filtering: exploiting item metadata in matrix factorization. *User Modeling and User-Adapted Interaction*, 29(2), pp.443-486.

16. Katarya, R., 2018. Movie recommender system with metaheuristic artificial bee. *Neural Computing and Applications*, 30(6), pp.1983-1990.

17. Guo, G., Zhang, J. and Yorke-Smith, N., 2015, February. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In Twenty-Ninth AAAI Conference on Artificial Intelligence.

18. Zhang, D., Hsu, C.H., Chen, M., Chen, Q., Xiong, N. and Lloret, J., 2013. Cold-start recommendation using bi-clustering and fusion for large-scale social recommender systems. *IEEE Transactions on Emerging Topics in Computing*, 2(2), pp.239-250.

19. Zhong, S., Zhang, W., Zhang, Q. and Lei, K., 2017, May. A trust networks recommender algorithm based on Latent Factor Model. In 2017 IEEE International Conference on Communications (ICC) (pp. 1-7). IEEE.

20. Thanh-Tai, H. and Thai-Nghe, N., 2017, November. A Semantic-Based Recommendation Approach for Cold-Start Problem. In International Conference on Future Data and Security Engineering (pp. 433-443). Springer, Cham.

21. Reshma, R., Ambikesh, G. and Thilagam, P.S., 2016, April. Alleviating data sparsity and cold start in recommender systems using social behaviour. In 2016 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 1-8). IEEE.

22. Revathy, V.R. and Pillai, A.S., 2019. A Proposed Architecture for Cold Start Recommender by Clustering Contextual Data and Social Network Data. In *Computing, Communication and Signal Processing* (pp. 323-331). Springer, Singapore.

23. Lee, M.R., Chen, T.T. and Cai, Y.S., 2016, August. Amalgamating social media data and

movie recommendation. In Pacific Rim Knowledge Acquisition Workshop (pp. 141-152). Springer, Cham.

24. Rosli, A.N., You, T., Ha, I., Chung, K.Y. and Jo, G.S., 2015. Alleviating the cold-start problem by incorporating movies facebook pages. *Cluster Computing*, 18(1), pp.187-197.

25. Ji, K. and Shen, H., 2016. Jointly modeling content, social network and ratings for explainable and cold-start recommendation. *Neurocomputing*, 218, pp.1-12.

26. Natarajan, S. and Moh, M., 2016, October. Recommending news based on hybrid user profile, popularity, trends, and location. In 2016 international conference on collaboration technologies and systems (CTS) (pp. 204-211). IEEE.

27. Moshfeghi, Y., Piwowarski, B. and Jose, J.M., 2011, July. Handling data sparsity in collaborative filtering using emotion and semantic based features. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval (pp. 625-634).

28. Ponnamp, L.T., Punyasamudram, S.D., Nallagulla, S.N. and Yellamati, S., 2016, February. Movie recommender system using item based collaborative filtering technique. In 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS) (pp. 1-5). IEEE.

29. Deldjoo, Y., Dacrema, M.F., Constantin, M.G., Eghbal-Zadeh, H., Cereda, S., Schedl, M., Ionescu, B. and Cremonesi, P., 2019. Movie genome: alleviating new item cold start in movie recommendation. *User Modeling and User-Adapted Interaction*, 29(2), pp.291-343.

30. Yi, P., Yang, C., Zhou, X. and Li, C., 2016, September. A movie cold-start recommendation method optimized similarity measure. In 2016 16th International Symposium on Communications and Information Technologies (ISCIT) (pp. 231-234). IEEE.

31. Pirasteh, P., Jung, J.J. and Hwang, D., 2014, April. Item-based collaborative filtering with attribute correlation: a case study on movie recommendation. In Asian conference on intelligent information and database systems (pp. 245-252). Springer, Cham.

32. Baccianella, S., Esuli, A. and Sebastiani, F., 2010, May. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec* (Vol. 10, No. 2010, pp. 2200-2204).

33. <http://github.com/word/emoji-emotion>, last accessed on Dec, 2020.