International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || PP.80-88

Analyzing Social Media Community Sentiment Score for Prediction of Success of Bollywood Movies

Sandeep Ranjan[1], Sumesh Sood[2]
[1]*(Department of Computer Science & Engineering, I.K. Gujral Punjab Technical University, Kapurthala, India)*
[2]*(Department of Computer Science & Engineering, I.K. Gujral Punjab Technical University, Kapurthala, India)*

Abstract: Social networks are developing as a real-time feedback system which can be used by product makers and service providers to fetch instant user acceptance and take corrective actions to increase their brand popularity. Like-minded users form communities when they post their opinions on to the social networks which affect the decision making of prospective customers. Often the largest community within a network represents the essence of the entire network in terms of sentiments. In the study presented, sentiment analysis of the largest community of the tweets about Bollywood movies released from 1st June 2017 and 17th, August 2017 was performed. The overall sentiment score divided by the number of members in the community when compared to Box-office collection of movies yields fairly good results.

Keywords: brand promotion, communities, online word of mouth, sentiment analysis, social networks

I. INTRODUCTION

The availability of freely downloadable data on the web servers and advanced search engines has greatly impacted user's information gathering capabilities [1]. The rapid growth of fast Internet access and low-cost devices and popular services such as Wikipedia, YouTube, blogs, online shopping websites, Facebook and Twitter a large amount of content is created by users with an exponential growth rate. One thing common in these services is that most of their content is generated by ordinary users and not by a few promoters. The general public generated content in social media is impacting individual user's day to day chores and decision making. The opinions and experiences in the social media content spread as online word of mouth and travel to both known and unknown users beyond geopolitical boundaries [2]. Studies reveal that a large proportion of users consult others or read their reviews before finalizing their purchase decisions and some are even ready to pay higher prices for higher rated service or brand.

Social media based word of mouth (WOM) is one of the best and most influential information source affecting purchase decision of consumers [3]. Such WOM affects pricing, demand and advertisement policies of brands. Online customers are more accessible and active and can provide cultural and marketing information that enables them to become a driving force in product design [4]. Social ties start developing among customers discussing a common issue and lead to the formation of communities. Social network communities discuss common topics and thus become a trusted source of information about a brand or product. Users participate in these communities to discuss their own experiences and concerns or they are interested in learning about other's engagement with the product. Customer loyalty can be gained by engaging customers through different relationship building activities such as blogs. Comparative analysis of different blogs reveals that reach of blogs to the masses is linked to effective blog writing [5]. Many corporations have come up with their blogs for brand building and customer engagement for a long-term human relationship [6].

Users of e-commerce websites trust the opinions posted about product manufacturers in an attempt to assess product and service standards [7]. Websites like www.ebay.com display star rating and a percentage positive feedback value for the product seller. Users can choose a seller on positive feedback satisfaction criteria which is the sum of past user experiences and opinions. Social network websites are evolving as the main source of information, communication and guide for purchase decisions about a wide variety of products and services[8]. Companies use social media as a platform to get access to new customers or engage their existing customers through a wide range of promotional strategies. Social media networks have become a means of fast feedback and platforms to develop models for consumer purchase behavior and societal trends. Semantic analysis on Airbnb Tweets to gain an insight of the service delivery reveals that user's posted opinions and experiences are of great value for service sector where the prospective customer's trust can be won by sharing this content [9]. Promotional strategies are very important in movie industry due to the experimental nature of movies and a large number of movies released over a short span of time. Such scenario is true for Bollywood Hindi movies where there are weekly movie releases. Professional critics generally give a star rating to movies which sometimes act as an initiator of word of mouth and customer engagement.

As the number of product launches is increasing, consumers are interested in the opinions of both professional critics and ordinary users [10]. Ordinary users of social media outnumber professional critics who are very less in number. The amount of content generated by ordinary users is very large compared to that generated by professional critics. In the research presented, the semantic analysis was performed on the tweets of both the ordinary users and professional critics for the Twitter handles of Bollywood movies released between 1st June 2017 and 17th August 2017. Recommendations of professionals are compared with those expressed by ordinary users and tested against the box office collection of the first week of movie launch.

II. LITERATURE REVIEW

Social media recommendation shave a broad scope both in industry and academia and many models are upcoming to analyze the role of different recommendation systems [11]. Max Ent and Naive Bayes algorithms were used to analyze the sentiments of Tweets for Hollywood movies to extract habits, opinions and preferences of customers [12]. The study classified Tweet sentiment as positive, negative and cognitive and also analyzed sentiments based on the country of origin. In addition to sentiment analysis, the genre of the movie, studies have carried out to predict genre of the movie from the Tweet datasets. Sentiment-based genre classification was done for Tamil movie Tweets to study the importance of user recommendation about Tamil movies [13]. A Tamil language dictionary "Agarthi" www.agarthi.com was used to derive the sentiment as the subject language was non-English. Movie recommendation and genre prediction system was developed based on Latent semantic analysis and singular value decomposition and it achieved 70% accuracy compared to IMDb genre data [14]. The study focused on developing a movie recommendation model based on semantic analysis of Tweets. Factors affecting the acceptance of recommendations of the social media followees were studied and the studies concluded that adaptive approaches give better results [15], [16].

Movie rating public datasets like Netflix and MovieLens have become obsolete with the advent of Tweet based sentiment analysis which is more user-centered[17].Analysis of reviews about 3,390 products listed on Amazon.com over 2 months period concludes that awareness about a product affects the purchase behavior of customers [18]. Movie success depends on a number of parameters like budget, a number of movies releasing in parallel, actors, directors and movie genre etc [19]. The study concludes that there is a correlation between the movie genre and performing actors and the model predicts the success of the concerned movie based on the actors-genre relationship. The online WOM activity on social media is at its peak before the movie release and in the first week of movie release [20]. For the general public, this is the crucial period when they seek recommendations and ratings of a new movie before making a decision to watch the movie. The analysis of WOM data from Yahoo Movie message board lowers the forecasting errors from 55% to 38% for weekly box office collection predictions.

Naïve Bayes and Support Vector Machine methods were compared for accuracy of sentiment classification of movie Tweets [21]. The study used a total of 1800 Tweets for training the model (600 positive Tweets, 600 negative Tweets and 600 neutral tweets). The accuracy of the model increases with the size of the dataset. First week box office collection of movie predicted by using clustering of features extracted from Tweets lowered the errors to 25% [22]. The study used the number of theater screens and the budget of the movie as a feature for clustering to develop a predictive model using both supervised and unsupervised learning for predicting movie per day revenue collection. Fuzzy inference system algorithm for data mining can be used for the sentiment analysis of movie Tweet datasets to classify the performance of a movie as a flop, average or a hit[23]. The method used two inputs, actor rating and different sentiment polarity score (negative, positive or neutral).Fuzzy inference system using hype created just before movie release gave fairly accurate result when compared to actual results of box office collection.

Communities play a major role in the spread of OWOM as they constitute like-minded users who discuss common issues [24]. There are retweets, sharing and linking of common interest post which adds on to the weight of the content being shared amongst community members. Due to the increasing popularity of social media as an inexpensive mode of communication, there is a motivation to detect communities which can help identifying participants and audience of the content. Community detection has helped law enforcement authorities to identify terrorist groups using various network properties [25]. Mapping Entropy Betweenness and Mapping Entropy centrality measures were used to identify Twitter users associated with terrorist activities. Identifying a terrorism-related user increases the probability of detecting other terrorism-related users. Top users in a community draft the topic of discussion while other common users mostly retweet or mention these tweets [26]. Clustering of Tweet datasets of Russian political protests resulted in bimodal community detection which includes Top users and common users based on centrality measures.

Detection of communities in social networks can be highly valuable to corporations for segregation of customers for target marketing [27]. Sentiment community detection applied to movie Tweet datasets reveal that ratings given by community members are consistent. Residents of two villages of the Indian state of Tamil Nadu were classified into communities based on their social relations with other villagers and the communities so detected were similar to their caste communities [28]. The study focuses on the interactions among similar entities leading to the concept of community structures in networks.

III. METHODOLOGY

The paper aims to develop a model to predict first week box office collection for Bollywood movies. Bollywood movies are normally released on Fridays, so Tweets for a particular movie were fetched from the release date to Thursday of following week. For predicting movie performance, two input components are used, the first component is the largest community's size and the second component is the total sentiment score of that community. The research analyzed the problems of community detection and sentiment mining from movie tweets for predicting the movie performance as a hit, average or flop.

3.1 Dataset Creation

Bollywood movies released between 1st June 2017 and 17th August, 2017 were selected for the study. The list of movie releases was tallied from sources like Wikipedia, ww.bollywoodhungama.com and www.bollywoodmdb.com. Tweets for each movie were fetched for 7 days starting from the movie release date. In all, data for 25 movies could be downloaded successfully, for some of the movies released in this period, more than one hashtags were popular on Twitter and hence those movies were not selected for the study. Twitter allows mining of public data using APIs and oauth authentication. Python and its libraries were used to fetch Tweets for the Twitter handles of Bollywood movies released in the study period. The data mining process and the dataset are restricted by the Twitter rate limit for the API being used. Tweets containing non-English characters were filtered out and duplicates were also removed from the datasets. Table 1 shows the number of Tweets and the number of distinct Tweets in each movie dataset.

Table 1. Number of distinct Tweets for each movie dataset

Movie Name	Release Date	Twitter handle	Distinct Tweets
Sweetie Weds NRI	02/06/17	sweetiewedsnri	138
Mirror Game	02/06/17	mirrorgamefilm	678
Hanuman da Damdadaar	02/06/17	hanumandadamdaar	1269
Flat 211	02/06/17	flat211	119
Dobaara	02/06/17	dobaara	2453
Dear Maya	02/06/17	dearmaya	1312
Behen Hogi Teri	09/06/17	behenhogiteri	1989
Raabta	09/06/17	raabta	4391
Love YouFamily	09/06/17	loveyoufamily	24
Bank Chor	16/06/17	bankchor	1747
Tubelight	23/06/17	tubelight	4657
Ek Haseena Thi Ek Deewana Tha	30/06/17	ehtedt	932
Mom	07/07/17	mommovie	59
Guest In London	07/07/17	guestinlondon	383
Jagga Jasoos	14/07/17	jaggajasoos	5105
Shab	14/07/17	shabthefilm	970
Lipstick Under My Burkha	21/07/17	lipstickundermyburkha	3360
Munna Michael	21/07/17	munnamichael	3720
Raag Desh	28/07/17	raagdesh	1362
Indu Sarkar	28/07/17	indusarkar	3123
Mubarakan	28/07/17	mubarakan	2463
Baarat Company	28/07/17	baaratcompany	54
Gurgaon	04/08/17	gurgaonthefilm	87
Jab Harry Met Sejal	04/08/17	jabharrymetsejal	148
Toilet Ek Prem Katha	11/08/17	toiletekpremkatha	6895

3.2 Betweenness Centrality

Social networks made up of a large number of nodes and multiple relations amongst the nodes can be depicted as graphs. Various centrality measures such as betweenness, closeness, degree, Eigenvector, load and page rank can be used to detect community structures in the graph representing the social network [26], [29]. The study used betweenness centrality measure to detect community structures in the movie datasets. The betweenness for each node is calculated as the number of edges connected to that node and is defined as
in terms of shortest paths that go through v.

1. An undirected, unweighted, connected graph G = < V, E >.
2. σ(s, t), the number of shortest paths existing between nodes s and t.
3. σ(s, t|v), the number of shortest paths existing between nodes s and t passing through node v.
4. The betweenness centrality of v, BC(v) is defined as:

$$CB\ (v) = \sum_{s \neq t \neq v} \frac{\sigma(s,t\,|v)}{\sigma(s,t\,)}$$

Betweenness centrality for all the nodes in a given graph involves computation of all the unweighted shortest paths between all node pairs of the graph. A breadth-first visit from a given source node v can be used to compute all the shortest paths from v in O(n + e) time complexity, where n and e are the number of nodes and edges of the graph. The study used Generalized Louvain method implemented in Python to detect communities [30]. A subgraph with a higher betweenness centrality value represents a well-connected, influential and concentrated information region [31]. Betweenness measure has been used in railway networks, freight networks and many other real-life networks to determine bottlenecks and other problems. A node with comparatively higher betweenness value than other nodes acts as a traffic checkpoint and can shut down or boost the network traffic [32].

Consider a graph made up of 7 nodes, A,B,C,D,E and F connected by edges as shown in figure 1. For each pair of nodes in this network, there exists a shortest path between them. The node that is located on maximum such shortest paths is the node with highest betweenness centrality.

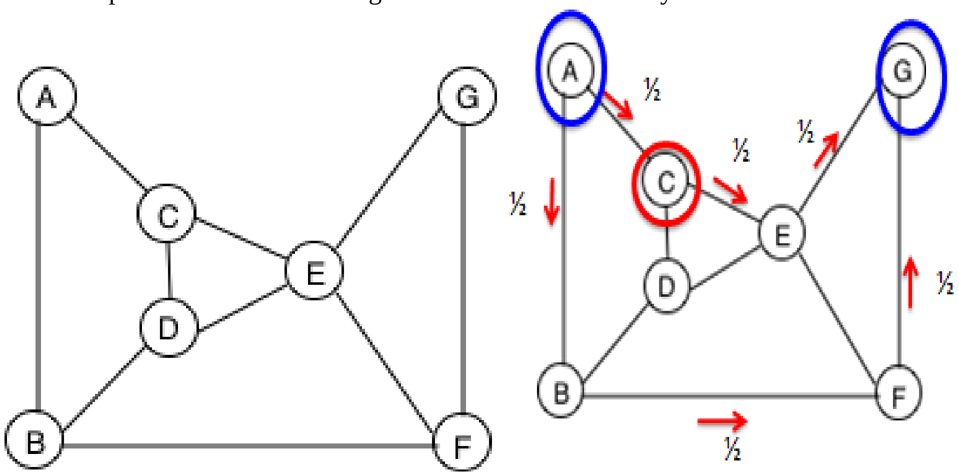

Fig. 1. Betweenness centrality of node C with respect to nodes A and G

Fig. 1 shows the basic procedure of calculating betweenness centrality in a network. Using a betweenness centrality measure, communities were identified in the movie datasets. In the social networks and particularly in the case of movie recommendation networks, the movie promoters (actors, producers and directors) and professional critics can initiate the conversations and have fair chances of the general public commenting on their content. This results in community creation and the growth of the community is dependent on the sentiment contained in the previously posted opinions.

Fig. 2 shows the communities in #sweetiewedsnri, #baaratcompany, #mommovie and #flat211 movie datasets. Nodes of each community have a distinct color and the largest community occupies a central position in the spiral graph representation of community structures. There are many communities in each network and the structure and number of communities depend on the relations amongst the nodes and the spread of online word of mouth as an opinion sharing phenomenon. Popular topics get more shares and like compared to less popular or unpopular ones. For each reply, retweet, share in the network, there is an edge created in the graph giving rise to communities of varying size and complexities. In figure 2, each graph has multiple communities and self-loops. From the point of view of movie popularity, popular movies create a bigger and better hype, get more attention and ultimately more recommendations (positive and negative).The largest communities in terms of number of nodes in them in each of the datasets are listed in Table 2.

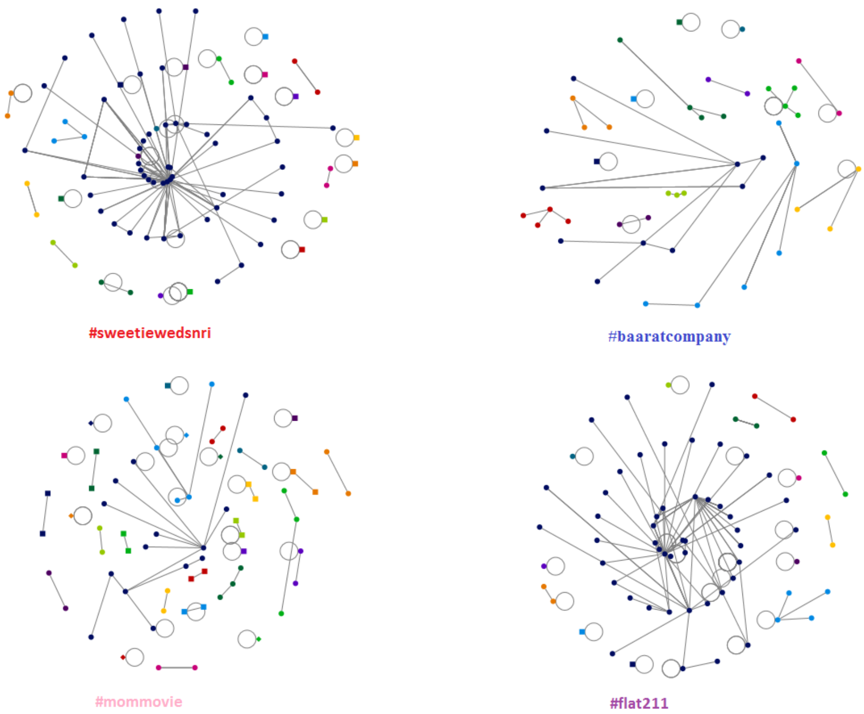

Fig. 2. Community structures of #sweetiewedsnri, #baaratcompany, #mommovie and #flat211 movie datasets

Table 2. Number of members in the largest community in each movie dataset

Movie Name	Twitter handle	No of nodes in the largest community
Sweetie Weds NRI	sweetiewedsnri	50
Mirror Game	mirrorgamefilm	262
Hanuman da Damdadaar	hanumandadamdaar	416
Flat 211	flat211	47
Dobaara	dobaara	617
Dear Maya	dearmaya	789
Behen Hogi Teri	behenhogiteri	686
Raabta	raabta	1343
Love YouFamily	loveyoufamily	7
Bank Chor	bankchor	652
Tubelight	tubelight	1879
Ek Haseena Thi Ek Deewana Tha	ehtedt	369
Mom	mommovie	14
Guest In London	guestinlondon	109
Jagga Jasoos	jaggajasoos	1657
Shab	shabthefilm	387
Lipstick Under My Burkha	lipstickundermyburkha	1478
Munna Michael	munnamichael	1471
Raag Desh	raagdesh	548
Indu Sarkar	indusarkar	1659

Mubarakan	mubarakan	1335
Baarat Company	baaratcompany	10
Gurgaon	gurgaonthefilm	24
Jab Harry Met Sejal	jabharrymetsejal	41
Toilet Ek Prem Katha	toiletekpremkatha	2451

3.3 Sentiment Analysis

Mere counting the Tweets and retweets related to a particular event does not give the essence of the event unless its sentiment analysis is carried out [33]. Opinions posted by social network site's users contain their sentiments which classify the content into positive, negative or neutral category. For brand owners it ie very important to know what their customers or prospective customers feel about the brand [34]. The present study focuses on the overall sentiment score of the largest community of the network. Once network communities have been identified, the task is to calculate the sentiment score of individual Tweet of the largest community. The semantic analysis was carried out on each of the movie hashtag's datasets to categorize the Tweets into five levels of polarity P, P+, N+, N and NEU.N+ and N indicate a negative polarity for a negative opinion in the content, whereas P and P+ indicate positive polarity. NEU polarity is generated when the polarity cannot be calculated.

TABLE 3- Polarity and weights

Polarity	N+	N	NEU	P	P+
Weight	-2	-1	0	1	2

Table 3 shows the polarities and their associated weights. NEU category Tweets have a neutral opinion about the entity or concept and hence have been assigned a zero weight. Negative opinion (N+ and N) Tweets have been assigned -2 and -1 weight. Positive opinion (P and P+) Tweets have +1 and +2 weights.

TABLE 4. Movie community overall sentiment score-community size ratio (N/M) (sorted by N/M score)

Movie Name	No of members in the largest community (M)	Community overall sentiment score (N)	N/M score
Flat 211	47	-124	-2.64
Sweetie Weds NRI	50	-89	-1.78
Baarat Company	10	-8	-0.80
Gurgaon	24	-13	-0.54
Dobaara	617	-244	-0.40
Raag Desh	548	-185	-0.34
Bank Chor	652	-147	-0.23
Jagga Jasoos	1657	-357	-0.22
Behen Hogi Teri	686	-125	-0.18
Mirror Game	262	-45	-0.17
Munna Michael	1471	-247	-0.17
Jab Harry Met Sejal	41	-6	-0.15
Guest In London	109	-14	-0.13
Ek Haseena Thi Ek Deewana Tha	369	-45	-0.12
Shab	387	-43	-0.11
Dear Maya	789	-59	-0.07
Indu Sarkar	1659	245	0.15
Mom	14	3	0.21
Raabta	1343	294	0.22
Tubelight	1879	477	0.25
Mubarakan	1335	347	0.26
Hanuman da Damdadaar	416	114	0.27

Love YouFamily	7	2	0.29
Lipstick Under My Burkha	1478	588	0.40
Toilet Ek Prem Katha	2451	1587	0.65

Box office collection is the best parameter to declare a movie as a flop, hit or average movie. The box office collection data was fetched from different websites such as www.bollywoodcat.com, www.bollywoodmdb.com, www.bollymoviereviewz.com, www.koimoi.com and www.jackace.in as there is not a single website which provides this data for all the Bollywood movies. The ratio of largest community's overall score to the number of nodes in that community can be used as a success parameter of the movie. This ratio has been compared with the Box office collection data which rates the movie success.

Table 5. Comparison of the community overall sentiment score-community size ratio (N/M) and box office collection

Movie Name	N/M score	Movie categorization on N/M score	Movie categorization on Box-office collection
Flat 211	-2.64	Flop	Flop
Sweetie Weds NRI	-1.78	Flop	Flop
Baarat Company	-0.80	Flop	Flop
Gurgaon	-0.54	Flop	Flop
Dobaara	-0.40	Flop	Flop
Raag Desh	-0.34	Flop	Flop
Bank Chor	-0.23	Flop	Flop
Jagga Jasoos	-0.22	Flop	Flop
Behen Hogi Teri	-0.18	Flop	Flop
Mirror Game	-0.17	Flop	Flop
Munna Michael	-0.17	Flop	Flop
Jab Harry Met Sejal	-0.15	Flop	Average
Guest In London	-0.13	Flop	Flop
Ek Haseena Thi Ek Deewana Tha	-0.12	Flop	Flop
Shab	-0.11	Flop	Flop
Dear Maya	-0.07	Flop	Flop
Indu Sarkar	0.15	Average	Flop
Mom	0.21	Average	Flop
Raabta	0.22	Average	Flop
Tubelight	0.25	Average	Average
Mubarakan	0.26	Average	Average
Hanuman da Damdadaar	0.27	Average	Flop
Love YouFamily	0.29	Average	Flop
Lipstick Under My Burkha	0.40	Hit	Hit
Toilet Ek Prem Katha	0.65	Hit	Hit

Table 5 data shows that the proposed model has successfully categorized 20 out of the total 25 movie datasets being studied. The model attained an accuracy level of 80%. It correctly categorized hit and flop movies with an exception of "Jab Harry Met Sejal" movie, but was not much efficient in the case of average movies.

IV. Conclusion

In social networks, popular contents (posts and Tweets) draw much public attention leading to a series of replies, shares and mentions leading to the evolution of communities in the network. The largest of the communities represents the essence of the entire network and its sentiment analysis can be used to represent the sentiment of the entire network. The success or failure of Bollywood movies is essentially dependent on the viewer recommendations which they share and gather from social networks like Twitter. The model presented in the study uses betweenness centrality to detect communities in the datasets and the sentiment analysis of the largest community when compared with actual Box-office collection gives 80% accuracy in categorizing movie as a hit, flop or an average movie.

References

[1] H. Rui and A. Whinston, "Information or attention? An empirical study of user contribution on Twitter," Inf. Syst. E-bus. Manag., vol. 10, no. 3, pp. 309–324, 2012.

[2] C. M. K. Cheung and M. K. O. Lee, "What drives consumers to spread electronic word of mouth in online consumer-opinion platforms," Decis. Support Syst., vol. 53, no. 1, pp. 218–225, 2012.

[3] B. A. Campbell, "American Economic Association Word-of-Mouth Communication and Percolation in Social Networks Author (s): Arthur Campbell Source : The American Economic Review ,Vol .103 , No . 6 (OCTOBER 2013), pp . 2466-2498 Published by : American Economic Associa," Am. Econ. Rev., vol. 103, no. 6, pp. 2466–2498, 2016.

[4] J. Brown, A. J. Broderick, and N. Lee, "Word of mouth communication within online communities: Conceptualizing the online social network," J. Interact. Mark., vol. 21, no. 3, pp. 2–20, 2007.

[5] S. Verma, "Online Customer Engagement through Blogs in India," J. Internet Commer., vol. 13, no. November, pp. 282–301, 2014.

[6] R. J. Brodie, L. D. Hollebeek, B. Jurić, and A. Ilić, "Customer engagement: Conceptual domain, fundamental propositions, and implications for research," J. Serv. Res., vol. 14, no. 3, pp. 252–271, 2011.

[7] B. Awrahman and B. Alatas, "Sentiment Analysis and Opinion Mining within Social Networks using Konstanz Information Miner," J. Telecommun. Electron. Comput. Eng., vol. 9, no. 1, pp. 15–22, 1843.

[8] and Q. L. N. M. Chinmay Pattnaik, "Innovation in social media strategy for movie success: a study of the Bollywood movie industry," Manag. Decis., vol. 56, no. 1, pp. 233–251, 2018.

[9] M. von Hoffen, M. Hagge, J. H. Betzing, and F. Chasin, "Leveraging social media to gain insights into service delivery: a study on Airbnb," Inf. Syst. E-bus. Manag., pp. 1–23, 2017.

[10] S. Moon, P. K. Bergey, and D. Iacobucci, "Dynamic Effects Among Movie Ratings, Movie Revenues, and Viewer Satisfaction," J. Mark., vol. 74, no. 1, pp. 108–121, 2010.

[11] J. Tang, X. Hu, and H. Liu, "Social recommendation: a review," Soc. Netw. Anal. Min., vol. 3, no. 4, pp. 1113–1133, 2013.

[12] U. R. Hodeghatta, "Sentiment analysis of Hollywood movies on Twitter," Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. - ASONAM '13, pp. 1401–1404, 2013.

[13] S. R. Nadana Ravishankar, "Corpus based Sentiment Classification of Tamil movie tweets using Syntactic patterns ARTICLE CORPUS BASED SENTIMENT CLASSIFICATION OF TAMIL MOVIE," vol. 8, no. June, pp. 172–178, 2017.

[14] S. Bansal, C. Gupta, and A. Arora, "User tweets based genre prediction and movie recommendation using LSI and SVD," 2016 9th Int. Conf. Contemp. Comput. IC3 2016, 2017.

[15] A. Tommasel and D. Godoy, "Learning and adapting user criteria for recommending followees in social networks," J. Assoc. Inf. Sci. Technol., 2017.

[16] M. Narayanan and A. K. Cherukuri, "A study and analysis of recommendation systems for location-based social network (LBSN) with big data," IIMB Manag. Rev., vol. 28, no. 1, pp. 25–30, 2016.

[17] S. Dooms, T. De Pessemier, and L. Martens, "MovieTweetings : a Movie Rating Dataset Collected From Twitter," Work. Crowdsourcing Hum. Comput. Recomm. Syst. CrowdRec RecSys, vol. 2013, no. January, pp. 43–44, 2013.

[18] J. Ren, W. Yeoh, M. Shan Ee, and A. Popovič, "Online consumer reviews and sales: Examining the chicken-egg relationships," J. Assoc. Inf. Sci. Technol., vol. 0, no. 0, 2017.

[19] J. Ahmad, P. Duraisamy, A. Yousef, and B. Buckles, "Movie Success Prediction Using Data Mining," in Computing, Communication and Networking Technologies (ICCCNT), 2017, pp. 1–4.

[20] Y. Liu, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," J. Mark., vol. 70, no. 3, pp. 74–89, 2006.

[21] A. Amolik, N. Jivane, M. Bhandari, and M. Venkatesan, "Twitter sentiment analysis of movie reviews using machine learning technique," Int. J. Eng. Technol., vol. 7, no. 6, pp. 2038–2044, 2016.

[22] S. Shim and M. Pourhomayoun, "Predicting Movie Market Revenue Using Social Media Data," 2017 IEEE Int. Conf. Inf. Reuse Integr., pp. 478–484, 2017.

[23] D. D. G. and B. Marakarkandy, "Using Twitter data to predict the performance of Bollywood movies," Ind. Manag. Data Syst., vol. 115, no. 9, pp. 1604–1621, 2015.

[24] S. Ranjan and S. Sood, "Identifying network communities in Bollywood Tweet dataset," Int. J. Adv. Res. Sci. Eng., pp. 503–512, 2017.

[25] I. Gialampoukidis, G. Kalpakis, T. Tsikrika, S. Papadopoulos, S. Vrochidis, and I. Kompatsiaris, "Detection of Terrorism-related Twitter Communities using Centrality Scores," Proc. 2nd Int. Work. Multimed. Forensics Secur. - MFSec '17, pp. 21–25, 2017.

[26] A. Semenov, I. Zakhlebin, A. Tolmach, and S. I. Nikolenko, "Pseudo-bimodal community detection in Twitter-based networks," Int. Congr. Ultra Mod. Telecommun. Control Syst. Work., vol. 2016–Decem, pp. 336–341, 2016.

[27] D. Wang, J. Li, K. Xu, and Y. Wu, "Sentiment community detection: exploring sentiments and relationships in social networks," Electron. Commer. Res., vol. 17, no. 1, pp. 103–132, 2017.

[28] C. De Bacco, E. A. Power, D. B. Larremore, and C. Moore, "Community detection, link prediction, and layer interdependence in multilayer networks," Phys. Rev. E, vol. 95, no. 4, pp. 1–10, 2017.

[29] Z. Yang, R. Algesheimer, and C. J. Tessone, "A comparative analysis of community detection algorithms on artificial networks," Sci. Rep., vol. 6, no. August, pp. 1–16, 2016.

[30] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized Louvain method for community detection in large networks," Int. Conf. Intell. Syst. Des. Appl. ISDA, pp. 88–93, 2011.

[31] T. Tsekeris, "Interregional trade network analysis for road freight transport in Greece," Transp. Res. Part E Logist. Transp. Rev., vol. 85, pp. 132–148, 2016.

[32] A. G. Nikolaev, R. Razib, and A. Kucheriya, "On efficient use of entropy centrality for social network analysis and community detection," Soc. Networks, vol. 40, pp. 154–162, 2015.

[33] A. Ceron, L. Curini, S. M. Iacus, and G. Porro, "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France," New Media Soc., vol. 16, no. 2, pp. 340–358, 2014.

[34] R. Sharma, V. Ahuja, and S. Alavi, "The Future Scope of Netnography and Social Network Analysis in the Field of Marketing," J. Internet Commer., vol. 2861, pp. 1–20, 2018.