INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

Detection of Fake Reviews using Linguistic Features

S.Uma Maheswari[1], D.R.Denslin Brabin[2]

Student, Department of Computer Science and Engineering, Madanapalle Institute of Technology & Science, Andhra Pradesh, India[1].

Associate Professor, Department of Computer Science and Engineering, Madanapalle Institute of Technology & Science, Andhra Pradesh, India[2].

ABSTRACT: Since the social media monitoring has been growing day by day, analysing of social media data plays an important role in knowing customer behaviour. So we are analysing social media data such as Twitter Tweets using sentiment analysis which checks the attitude of user review on movies, shopping, etc.Since writing fake reviews / frauds comes with significant gains, there has been a huge increase in fake spam views on online review websites.This paper proposes a method to detect fake reviews that develops a combined dictionary based on social media keywords and online review and also finds hidden relationship pattern from these keywords. A good review of the target item can attract more customers and increase sales; Poor reviews of the target item may result in lower demand and decreased sales. This false / fraudulent review was deliberately written to mislead potential customers in order to induce / deceive or defile their prominence. Our work aims to identify whether the review is false or factual.

KEYWORDS: Classification, E-Commerce, Fake Review, Linguistic Features, Sentiment Analysis

I. INTRODUCTION

Nowadays, Social media is becoming more and more popular since mobile devices can access social network easily from anywhere. Therefore, Social media is becoming an important topic for research in many fields. As number of people using social network are growing day by day, to communicate with their peers so that they can share their personal feeling every day and views are created on large scale. Social media monitoring or tracking is most important topic in today's current scenario. In today many companies have been using Social Media Marketing (SMM) to advertise their products or brands, so it becomes essential for them that they can be able to calculate the success and usefulness of each product [1]. For Constructing a social media monitoring, various tool has been required which involves two components: one to evaluate how many users of their brand are attracted due to their promotion and second to find out what people thinks about the particular brand. To evaluate the opinion of the users is not as easy as it seems to all users. For evaluating their attitude may require to perform Sentiment Analysis, which is defined as to identify the polarity of customer behavior, the subjective and the emotions of particular document or sentence.

To perform sentiment analysis, we needMachine Learning (ML) and Natural Language Processing (NLP) methods and this is place where most of the developers facing difficulty when they are trying to form their own tools. Over the recent years, an emerging interest has been occurred in supporting social media analysis for advertising, opinion analysis and understanding community cohesion. Social media data adapts to many of the classifications attributed for "big-data" – i.e. volume, velocity and variety. Analysis of Social media needs to be undertaken over large volumes of data in an efficient and timely manner. Analyzing the media content has been centralized, due to the key role that the social media plays in modelling public opinion. This type of analysis typically on the preliminary coding of the text being examined, a step that involves reading and annotating the text and that limits the sizes of the data that can be analyzed. With the development of Web, more and more people are connecting to the Internet and becoming information producers instead of only information consumers in the past, resulting to the serious problem of information overloading. There is much personal information in online textual reviews, which plays a very important role on decision processes. For example, the customer will decide what to buy if he or she sees valuable reviews posted by others, especially user's trusted friends. People believe reviews and reviewers will do help to the rating prediction based on the idea that high-star ratings may greatly be attached with good reviews.

Hence, how to mine reviews and the relation between reviewers in social networks has become an important issue in web mining, machine learning and natural language processing. It focusses on the rating prediction task.Fig.1

shows an example of positivereviews and negative reviews on website. From Fig.1, thereare many positive words in a 5-star review, such as "great", and "lovely". But in a 2-star review we find negative words,such as "expensive", and "poor". That means a good reviewreflects a high star-level and a bad review reflects a low-level.When we know the advantages and disadvantages from thetwo kinds of reviews, we can easilymake a decision. Detection of positive and negative reviews involve sentiment analysis and text analysis.
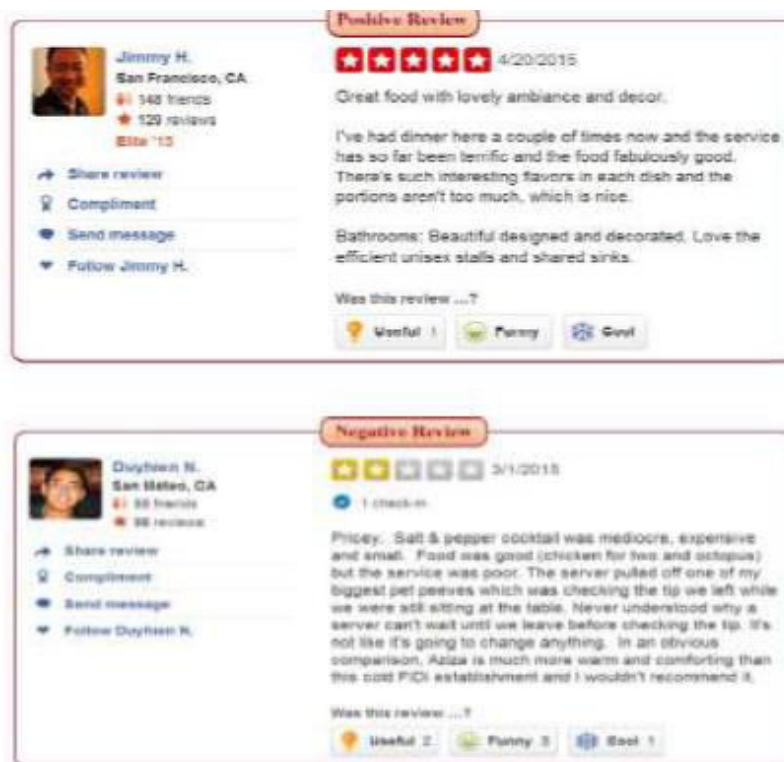

Fig. 1. An Example of Positive Review and Negative review on websites

Sentiment analysis refers to the use of natural language processing to identify and extract one-sided information in source materials or simply it refers to the process of detecting the polarity of the text. It also referred as opinion mining, as it derives the opinion, or the attitude of a user. A common approach of using this is how people think about a particular topic. Sentiment analysis helps in determining the thoughts of a speaker or a writer with respect to some subject matter or the overall contextual polarity of a document. The attitude may be his or her decision or estimate, the emotional state of the user while writing [2].Today, Sentiment analysis plays an important role where various machine learning technique is used in determining the sentiment of very huge amounts of text or speech. Various application tasks include such as determining how someone is excited for an upcoming movie, correlates different views for a political party with people's positive attitude towards vote for that party, or by converting written hotel reviews into 5-star based on scaling across categories like 'quality of food', 'services', 'living room' and 'facilities' provided. As there is huge amount of information is shared on social media, forums, blogs, newspaper etc. it is easy to see why there is a need for sentiment analysis as there is much information to process manually which is not possible in today's time.

Text analysis involves three processes: Data acquisition, Preprocessing and Data mining. In this data acquisition, data are gathered from different relevant sources such as web crawling, twitter tweets, online review, newsfeeds, document scanning etc.Preprocessingis used to remove noisy, inconsistent and incomplete data. For doing the classification, text preprocessing and feature extraction are the preliminary phases.Preprocessing involves tokenization or segmentation and removal of stop words. Researchers have shown that by removing stop words from the file, you can get the benefit of reduced index size without much affecting the accuracy of a user.With the help of eliminating stop words from the index, the index size can be reduced to about 33% for a word level index. While assessing the content of natural language processing, meaning of word can be conveyed more clearly by removing the functional word. Data Mining involves applying different mining techniques to derive usefulness about stored information. Different mining approaches are classification, clustering, statistical analysis, natural language processing etc.

In text analytics, mainly classification technique is used. Classification is a supervised learning method that helps in assigning a class label to an unclassified tuple according to an already classified instance set. Data classifying and identifying is all about to tag the data so it can be created quickly and efficiently. But various organizations can gain from re-transforming their information, which helps in order to cut storage and backup costs, with increasing the speed of data searches. Classification can help an organization to meet authorized and regulatory requirements to retrieve specific information within a specific time period, and this is most important factor behind implementing various data classification technology. Analytical results provide valuable things from text mining so that it can provide information that helps in improving decision and processes. It includes following ways such as sentiment analysis, document imaging, fraud analysis etc.

II. LITERATURE SURVEY

Reasonable researches were carried out in the field of review analysis. Hennig-Thurau et. al. [3] states that customer comments articulated via the Internet are available to a large number of other customer's, and therefore can be expected to have a significant impact on the success of goods and services. Consumer buying and communication behaviour are tested in a large-scale empirical study. The results illustrate that consumers read online articulations mainly to save decision-making time and make better buying decisions. Structural equation modelling shows that their motives for retrieving online articulations strongly influence their behaviour.Chevalier et. al. [4] hypothesized that buyerssuspect that many reviewers are authors or other biased parties. They found marginal (negative) impact of 1-star reviews is greater than the (positive) impact of 5-star reviews. The results suggest that new forms of customer communication on the Internet have an important impact on customer behaviour. Work on sentiment analysis found using a formal approach in [5]. It presents a method to detect sentiment of newspaper headlines, in fact partially using the same grammar formalism that later will be presented and used in this work, however without the combinatorial logic approach. This focuses on some specific problems arising with analysing newspaper headlines, e.g. such as headline texts often do not constitute a complete sentence, etc. However,it also presents more general methods, including a method for building ahighly covering map from words topolarities based on a small set of positive and negative seedwords. This method has been adopted here, solves the assignment of polarity values on the lexical levelquite elegantly, and is very loosely coupled to the domain.

Wang et. al. [6] proposed a method for rating the reviews instead of binary classification. Review rating is more useful than the two-way review classificationfor many decision processes. In the review rating, reviewers are influenced not only by their own subjective feelings, but also by others' rating to the same product. It considers the content of reviews, subjective factors of reviewers and the impact of other people in the social relations of reviewers. Luo et. al. [7] stated an approach to predict the ratings for the reviews that are not rated. They proposed a model based on Latent Dirichlet Allocation (LDA) with indirect supervision that uses the quad-tuples of (head, modifier, rating, entity) to explicitly show the associations between modifiers and ratings. Another method for rating prediction by exploring service reputation is projected in [8]. How to mine valuable information from these reviews and make an accurate recommendation is very importent in review analysis. In this method, social user's reviews are analysed and each user's sentiment score on items/services is calculated. Then the service reputation factor is used in the recommender system to make an accurate rating prediction

Electronic word-of-mouth (eWOM) is a form of communication, defined as a:"statement made by potential, actual, or former customersabout a product or company, which is made available to amultitude of people and institutions via the Internet".eWOM may be less personal in that it is not face-to-face (ormaybe just personal in a different way than in the past), butit is more powerful because it is immediate, has asignificant reach, is credible by being in print, and isaccessible by others.Different types of data are generated from different Socialmedia groups that need to be organized and to monitorpeople's attitude towards products, gadgets, movie reviewetc. This database is collected from different social mediasites for example Twitter, Face book, Online review,shopping sites etc. Text analytics and Sentiment analysiscan help to develop valuable business insights from textbased contents that may be in the form of word documents,tweets, comments and news that related to Social media.The foremost reason of sentiment analysis is so complex that word often takes different meanings and are associatedwith different emotions depending on the domain in whichthey are being used

Matsumoto et. al. [9] proposed an approach for movie recommendation based on the prediction of movie linksaccording to the local and global structures of a network. Personalized video recommendation by using the contents (e.g., text, audio, motion, and so on) of the videos is explained in [10]. Here a general framework is proposed that includes content feature with user-video interactions, named as collaborative embedding regression model, to make effective video recommendation in all scenarios.It is practically impossible to distinguish between fake and genuine reviews by only reading their content. However, the good results can be obtained by using linguistic features on a

domain specific dataset (i.e., a Yelp's dataset containing only New York Japanese restaurants). It is possible to use Natural Language Processing techniques to extract simple features from the text, and to use as features some statistics and some sentiment estimations connected to the use of the words.

Three primary methods are available and used more for text classification because of their relative popularity and success in prediction of sentiments:

1. Naive Bayes:This works on the assumption of conditional independence and despite this oversimplified assumption, Naive Bayes performs well in many complex real-world problems. Naive Bayes classifier is superior in terms of CPU and memory consumption.
2. Support Vector Machines (SVM):It also provides a robust approach to build text classifiers and was picked because of its ability to handle high dimensional input space. When learning text classifiers, many (more than 10000) features can been countered. Since SVMs use over fitting protection, which does not necessarily depend on the number of features, they have the potential to handle these large feature spaces.
3. Maximum Entropy**:** It is based on conditional independence assumption, which uses maximum entropy that does not assume conditional independence. Although it takes more time than Naïve Bayes to train the model, this method has proven to be useful in cases where we do not know anything about the prior distribution

III. PROPOSED SYSTEM

As we briefly introduced in Section II, many and different features that have been considered so far in the review site context to identify fake reviews. In some cases, features belonging to different classes have been considered separately by distinct approaches. In other cases, the employed features constitute a subset of the entire set of features that could be taken into account; furthermore, new additional features can be proposed and analyzed to tackle open issues not yet considered, for example the detection of singleton fake reviews. For these reasons, in this section we provide a global overview of the various features that can be employed to detect fake reviews. Both significant features taken from the literature and new features proposed in this paper are considered to detect fake reviews. The most effective approaches are in general supervised and consider review-centric and reviewer-centric features.

Reviewcentric features that have been considered, is constituted by those related to a review. They can beextracted both from the text constituting the review, i.e., textual features, and from meta-data connected to a review, i.e., metadata features. In every review site, the time information regarding the publication of the review, and the rating (within somenumerical interval) about the reviewed business are metadata, are always provided. In addition to metadata features,those connected to the cardinality of the reviews written by a given user must be carefully studied. In fact, a large part of reviewsare singletons, i.e., there is only one review written by a given reviewer in a certain period of time (this means that in the useraccount there is only one review at the time of the analysis). For this kind of reviews, specific features must be designed. Many of the features are based on some statisticsover several reviews written by the same reviewer. In the case of singletons, these features loose their relevance in assessingcredibility.Therefore, the definition of suitable features that are effective for detecting singleton fake reviews also becomes crucial.
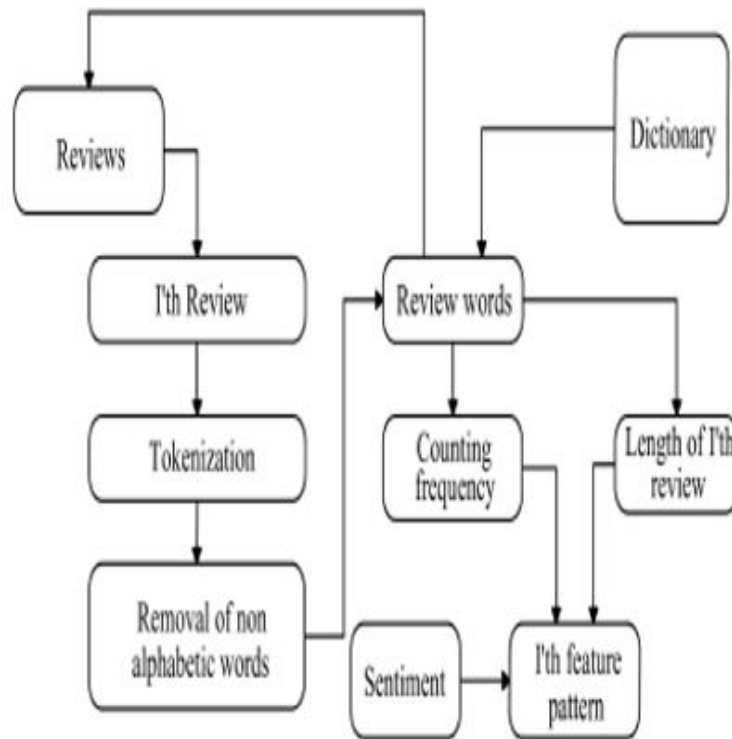
Fig.2. Review Analysis Process

The review analysis process of our proposed system is shown in Fig 2. Preprocessing is the important step in our proposed system as like other review analysis system. For doing the classification, text preprocessing and feature extraction is a preliminary phase. Preprocessing involves there steps: Word parsing and tokenization, Removal of stop words and Stemming. In word parsing and tokenization phase, each user review splits into words of any natural processing language. As movie review contains block of character which are referred to as token. Stop words are the words that contain little information, so needed to be removed. As by removing them, performance increases. Here, we made a list of around 320 words and created a text file for it. So, at the time of preprocessing we have concluded that so all the stop words are removed from our dataset i.e. filtered. Stemmingis defined as a process to reduce the derived words to their original word stem. For example, "talked", "talking", "talks" as based on the root word "talk". We have used Snowball stemmer to reduce the derived word to their origin.

Classification is a supervised learning method that helps in assigning a class label to an unclassified tuple according to an already classified instance set [11]. Here, naïve bayes multinomial classifier has been used. Quality measure will be considered on the basis of percentage of correctly classified instances. For the validation phase, we use 10- fold cross validation method. Naïve bayes multinomial helps in generating dictionary and frequent set. It counts the occurrences of words in whole dataset and forms a dictionary of some most frequently occurring words.Here, we analyze the dataset based on accuracy given by naïve bayes multinomial classifier as shown in Fig. 3. Online review dataset accuracy around 94.968% and for twitter its around 82.695%. Results show that we get better accuracy for online review as compared to twitter tweets as online review are more clear and in detail compare to twitter tweets.Combined dictionary of words of twitter tweets and online review are formed based on probability of each word as we get by classification algorithm.
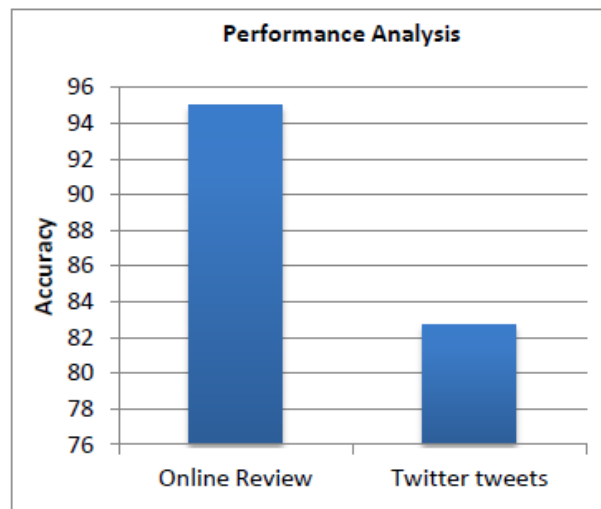
Fig. 3. Performance analysis based on accuracy

In this proposed system, hadoop open source data mining tool has been used to perform sentiment classification on movie review dataset. The goal is to classify dataset into positive and negative and form the combined dictionary of Twitter dataset and online review dataset. Two dataset were collected firstly, from Twitter tweets and secondly, from Online review Dataset. The online review dataset consists of around 800 user's review archived on the IMDB (Internet Movie Database) portal. And for, Twitter dataset around 1000 review were collected and each review were formatted according to .arff file where review text and class label are only two attributes. Class label represent the overall user opinion. Simple rules for scaling the user reviews are generated. For dataset, a user rating greater than 6 is considered as positive, between 4 to 6 considered as neutral and less than 4 considered as negative. Dataset can be analysed by using the weka tool. The hidden relationship has to be extracted from this type of database using different mining approaches in Weka tool. Dictionary building for detailed sentiment analysisimplies making an initial list of adjectives and nouns which are normally used when describing a specific movie review. Phrases and terms are extracted from this relational dataset and their meaning has been added to dictionary for next generation analysis. In tweets, informal and shortcuts has been used for explaining terms or views and this is done with the help of sentiments analysis is not an easy process. To reduce this, data mining approaches has been used for extraction of features from these datasets. Word embedding methods have been extensively recommended for sentiment classification [12]. Word2Vec and GloVe are the most precise and functioning word embedding methods which can translate words into meaningful vectors in the current era.

Similar reviews are identified by text analysis, keyword matching and meta data matching techniques. Afterwards the reviews are analyzed for positive and negative reviews. Then analysis is extended for multiple review from same user-ID and multiple reviews from same IP address. Based on all these analysis, reviews are rated or ranked. Fake reviews are identified using the rating. Reviews with high rate are identified as fake reviews and reports are generated.

IV. RESULTS

The proposed fake review detection system is implemented as online system using JSP. The system can be applied to different types of review analysis system such as shopping, product, movie, hotel etc. Here we are showing an example of movie review analysis system. Fig. 4 shows the report of some sample movie reviews and Fig. 5 shows the report on movie rating. After applying all the analysis on the reviews, fake reviews are identified. Fig. 6 shows report of some sample fake reviews.

Id	User Name	Movie Name	Review Details	Date and Time	Region
3	Omkar	The Villain	nice	09/09/2019 17:35:25	vijayanagar
4	Omkar	The Villain	villain is good	09/09/2019 13:03:00	vijayanagar
5	Omkar	The Villain	villain is bad	09/09/2019 13:03:30	vijayanagar
6	Omkar	The Villain	villain is bad	09/09/2019 13:04:02	vijayanagar
7	Manjunath	Onti	It is bad movie	09/09/2019 17:16:29	Malleshwaram
8	Manjunath	Onti	It is good movie and beautiful	09/09/2019 17:16:48	Malleshwaram
9	Manjunath	Robot2	It is good movie	09/09/2019 16:12:50	Malleshwaram
10	Manjunath	Robot2	It is totally animation based project.	09/09/2019 16:36:14	Malleshwaram
11	Ramesh	Mission_Mangal	It is best movie	09/09/2019 18:17:16	Malleshwaram
12	Ramesh	Mission_Mangal	It is bad movie	09/09/2019 18:17:27	Malleshwaram
13	Manjunath	Mission_Mangal	It is good movie	09/09/2019 18:18:46	Malleshwaram
14	Manjunath	Mission_Mangal	It is beautiful film	09/09/2019 18:19:53	Malleshwaram

Fig. 4. Report of sample reviews

Id	User Name	Rate	Movie Name	Date and Time	Region
1	Manjunath	4	Robot2	09/09/2019 17:14:54	Malleshwaram
2	Manjunath	3	Rjakumara	09/09/2019 17:26:35	Malleshwaram
3	Manjunath	1	Kaala	09/09/2019 17:31:43	Malleshwaram
4	omkar	8	Kaala	09/09/2019 17:39:35	vijaynagar
5	Ramesh	7	Mission_Mangal	09/09/2019 18:17:01	Malleshwaram
6	Ramesh	3	Mission_Mangal	09/09/2019 18:18:04	Malleshwaram
7	omkar	1	Mission_Mangal	09/09/2019 18:48:22	vijaynagar
8	sai	3	The Villain	13/02/2020 11:02:08	kd

Fig. 5. Report of review rating

Id	User Name	Movie Name
2	Manjunath	Kaala
5	Ramesh	Mission_Mangal
2	Manjunath	Onti
2	omkar	Rjakumara
2	Manjunath	Robot2
8	Omkar	The Villain

Fig. 6. Report of fake reviews

V. CONCLUSION

Determining and categorizing reviews as false or factual is an important and challenging issue. In this paper, we proposed an online fake review detection system that have used language features such as presence of unigram, frequency of unigram, presence of bigram, frequency of bigram and length of reviews to build the model and detect fake reviews. For obtaining false reviews it requires both linguistic and behavioral features. The proposed model focuses more on obtaining confidential reviews using supervised learning on language features. The same model can be extended in future by combining behavioral and linguistic features using supervised or unsupervised learning methods.

REFERENCES

1. M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang, "Social contextual recommendation",in proceedings of the 21st ACM international conference on Information and knowledge management, pp. 45–54, 2012.
2. Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis", in proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, pp. 83–92, 2014.
3. T.Hennig-Thurau, G. Walsh, and G. Walsh, "Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet", International journal of electronic commerce, vol. 8, no. 2, pp. 51-74, 2003.
4. J. A. Chevalier, and D. Mayzlin,"The effect of word of mouth on sales: Online book reviews", Journal of marketing research, vol. 43, no. 3, pp. 345-354, 2006.

5. F.Simancík, and M. Lee, "A CCG-based system for valence shifting for sentiment analysis", Research in computing science, vol. 41, pp. 99–108, 2009.
6. B. Wang, Y. Min, Y. Huang, X. Li, and F. Wu, "Review rating prediction based on the content and weighting strong social relation of reviewers", in Proceedings of the ACM international workshop of mining unstructured big data using natural language processing, pp. 23-30, 2013.
7. W. Luo, F. Zhuang, X. Cheng, Q. He, and Z. Shi, "Ratable aspects over sentiments: Predicting ratings for unrated reviews", in Proceedings of the IEEE international conference on data mining, pp. 380-389, 2014.
8. X. Lei, and X. Qian, "Rating prediction via exploring service reputation",in Proceedings of the IEEE 17th international workshop on multimedia signal processing, pp. 1-6, 2015.
9. Y. Matsumoto, R. Harakawa, T. Ogawa, and M. Haseyama, "Music movie recommendation based on link prediction considering local and global structures of a network", IEEE Access, vol. 7, pp. 104155–104167, 2019.
10. X. Du, H. Yin, L. Chen, Y. Wang, Y. Yang, and X. Zhou, "Personalized video recommendation using rich contents from videos",IEEE transactions on knowledge and data engineering, vol. 32, no. 3, pp. 492-505, 2020.
11. J. K. Rout, A. Dalmia, and K.-K. R. Choo, "Revisiting semi-supervised learning for online deceptive review detection", IEEE Access, vol. 5, pp. 1319–1327, 2017.
12. S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings" Expert System and Applications, vol. 117, pp. 139–147, 2020.

INTERNATIONAL JOURNAL
OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING