A Novel Adaptable Approach for Sentiment Analysis

Aishwarya R, Ashwatha C, Deepthi A, Beschi Raja J

Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, Tamil Nadu, India

ABSTRACT

The internet has provided many novel ways for people to express their ideas and views about different topics, ideas and trends. The contents generated by the users which are present on various mediums like internet blogs, discussion forums, and groups paves a strong base for decision making in diverse fields such as digital advertising, election polls, scientific predictions, market surveys and business zones etc. Sentiment analysis is the process of mining the sentiments from the data that are available in online platforms and categorizing the opinion towards a particular entity that falls on three different categories which are positive, neutral and negative. In this paper, the problem of sentiment classification of election dataset in twitter has been addressed. This paper summarizes the ensemble method, the best way to achieve classification. And also about the ada boosting algorithm and artificial neural networks by which the optimized prediction accuracy is achieved.

Keywords : Sentiment Analysis, Twitter, Opinion Mining, Social Media.

I. INTRODUCTION

Social computing is an emerging and an innovative trend for analyzing and modelling many activities that are taking place in many social platforms. Many interesting, intellectual and interactive applications are produced to reach efficient outcomes. A diverse range of social media sites are available which provides the individuals to express their opinions about a particular event or an issue. According to many surveys, millions of people use the social sites to express their thoughts. People may convey anything on such sites like they do comments on the any products or posts. They can share it with the others. Moreover, it's really clear that they provide a greater opportunity to the business people by giving a platform in where they can connect with their customers by which they can advertise or speak directly with their customers.

On the other side, consumers holds the power when it comes to what consumers wants to see and how they are responding to it. By this, the success & failure of the company is shared publicly which ends up with word of mouth. However, the social networking methods changes the behavior and decisions of the consumers, for example, mentioning that 90% of internet users are getting influenced on their purchase and decision by customer's review. So that, if organization could catch up faster on what their customers are thinking, it is more beneficial to organize and react in time by which they can come up with a good strategy to raise above their competitors.

II. RELATED WORK

Many studies have focused on the analysis of the social media contents, especially in case of some big
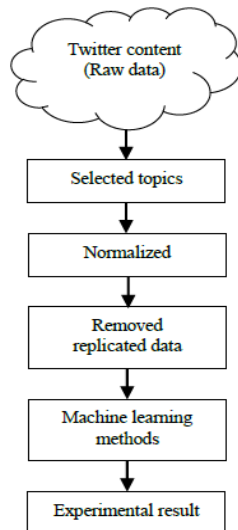
events which attract greater attentions like presidential elections[3]. Social media can be said as an important platform for both the sides, candidates who needs to share about about their programs and to be in direct touch with the people, and the voters where they can express themselves about each of the candidates[1]. The in depth usage of social media had attracted greater attentions in educational research and a larger number of contributions had been conducted to do follow those kind of events.

Many authors had done research for the logical approach for mining the sentiments that are shared on various social media sites[1]. They had done analysis on the sentiments of various texts by using systems like combinatory category grammars, lexicons, annotation, and semantic text networks. The base techniques for sentiment classification and the base methods for collecting the data are presented[2]. The desired accuracy for classification process with a selected vector has been verified for many electronic product categories using many different classifiers like Maximum Entropy method, Naive Baye's method, Support Vector Machines, and many Ensemble classifier methods. Authors had also introduced a thing called hybrid method which is a combined usage of the sentiment lexicons with a machine learning classifier for the detection of subjective texts in the many consumer-products sub domains[2]. Authors had proposed a diverse batch of machine learning techniques with semantic analysis for classifying the sentence and also the reviews of various products based on the twitter data for better accuracy. Authors had clearly examined the performance of the classifiers like as Naïve Bayes, SVM, SMO, and Random Forest for the better classification of Twitter data.Authors had clearly presented the technique for normalizing the irrelevant tweets and classifying them according to the polarity like positive,negative or neutral. Moreover, they had been employing a mixture model technique to generate various sentimental words[2].

The words that are generated are used in later as the feature indicators on the classification model. Authors have invented a novel approach method to predict the sentiments about the stocks by various monetary boards and have done performing an automatic prediction for the stock market by the usage web sentiments. Authors have examined the performance of sentiment analysis in e-learning domain, automatic sentiment classifier to classify reviews of Brazilian TV show to predict the positive or negative category and attained nearly 90% of accuracy[3]. They had deployed a system for extracting the Tweets and classifying them using the domain oriented seed based enrichment technique for the purpose of information loss in the knowledge of the domain. Authors had developed an optimized sentiment classification technique which can detect the anomalies from the Twitter data in addition with the classification.

III. METHODS FOR SENTIMENT ANALYSIS

The process of sentimental analysis in the field of Twitter data is a growing field which needs greater pack of attention. There exists certain procedures for carrying out the process of sentimental analysis on the Twitter dataset. Firstly, the Twitter data which had been collected is pre-processed in order to perform the data cleaning. Next, the evidential features are needed to be extracted from the clean text, by applying any feature selection methods available. Thirdly, a little portion of the data is labeled as either positive or negative or neutral Tweets manually, to prepare a proper training set. Finally, the features are extracted and the labeled training set is provided as an input to the built classifier for classifying the remaining data i.e. test data set. Each of the steps is described thoroughly in the following subsections.

```
        Twitter content
         (Raw data)

              ↓

        Selected topics

              ↓

          Normalized

              ↓

           Removed
        replicated data

              ↓

       Machine learning
           methods

              ↓

      Experimental result
```

A. Data Sources

Selecting the data source for conducting the sentiment analysis has a significant role[3]. The data sources in the social media sites are broadly classified into three general categories: blogs, micro-blog sites, and review site. Among those categories, a micro-blogging one like as Twitter has attained higher popularity because to its limited strength of the content and the data that are available publically[4]. Based on many of the statistics of the Twitter growth rate, it's provident to use Twitter as the data source for the sentimental analysis.

Growth Rate Statistics

Eventually 6,000 tweets are done on the Twitter on a second basis. It resembles to 3,50,000 tweets sent per minute and 500 million tweets per day. That makes it around 200 billion tweets per year[4]. In Twitter's history, the number of Tweets had been increased from 5,000 per day to 500,000,000 tweets per day that is approximately a six orders of magnitude].What's more, 340 million tweets for every day six years after the development of Twitter for example on March 21, 2012 This measurements finish up the utilization of Twitter for our examination reason[5].

Twitter Studies

As per the recent survey, the studies carry out on Twitter datasets are in the field of health, stock markets, politics, advertising, sports etc. Analysis techniques used in these studies are many and they include qualitative content analysis, network analysis, graph analysis, linguistic analysis or psycholinguistic analysis, word clouds and histograms[6]. In addition to that, the Twitter has been toply voted as the most effective source in the studies such as community and influence detections, topic discoveries, marketing and business predictions.

Tweets

The messages or comments that are posted on the Twitter is called a Tweet, which is limited to 140 characters or less.Tweets typically composes of any of the followings text, links, emoticons, and images[7]. Upon these components the mining technique is applied to classify the text, links, images, emojis or emoticons and even videos. The Tweets contains three documentations including hashtags (#), retweets (RT) and record Id (@).

B. Twitter Data Collection Methods

There are three ways to collect the Tweets for research research purposes.Information stores, for example, UCI, Friendster, Kdnuggets, and SNAP.

API- Twitter allows two types of APIs such as search API and stream API[8].

C. Preprocessing of Data

Mining of any data is a challenging role. The collected data which are collected is a raw data. So, in order to apply the classifiers, it's really essential to pre-process the raw data that are collected[8]. The pre-processing task involves uniform casing which means the conversion of upper to lower case or vice versa, removing the hashtags and other notations like @, RT, emojis, URLs, stop words, conversion or removal of slang words and compression of elongated

words[9]. The following steps show the pre-processing procedure.

Removing the Twitter notations like hashtags (#), retweets (RT), and account Id (@).

Removing the URLs, hyperlinks and emoticon. Every non letter data and symbols are removed as we deal with only text data.

Removal of the stop words like are, is, am etc, it is intended to remove them to compress the dataset[9].

Compressing of the elongated words such as youuu into you. Decompressing the slang words such as ni8, fi8. Generally slang words contain the extreme level of sentiments[10]. So it is necessary to decompress them.

D. Feature Extraction

The dataset that has been preprocessed has various properties. In feature extraction, the different aspects such as adjectives, verbs and nouns are extracted and after extracting them their aspects are identified as positive or negative to detect the polarity of the whole sentence. Followings are the widely used Feature Extraction methods

Terms Frequency and Term Presence: These features denote individual and distinct words and their occurrence counts[11].

The negative Phrases, for example, the presence of negative words changes the whole meaning or an orientation of the opinion. So it is proven to take the negative word orientation in account[11].

The Parts Of Speech (POS) like finding the nouns, verbs, adjectives etc. as they are an important metrics of opinions.

E. Sentiment Classification Techniques

Typically two techniques are used to identify the sentiment of any text. One is knowledge based technique and an another one is machine learning techniques[12]. Knowledge based technique is nothing but the Lexicon based technique that mainly

focuses on deriving the opinions only based on the lexicons from the text and then the polarity of those lexicons are identified. Lexicons are nothing but the group of known and precompiled sentimental terms[13]. This technique is further classified into Dictionary-based and Corpus based approach. In the Dictionary-based model, the opinion oriented words are found, and then the dictionary is examined to collect respective synonyms and antonyms[14]. Whereas in the Corpus based technique, a list of opinion words are created and then based on their context specifications and orientations, additional related words are found in a vast corpus[14]. To perform a lexicon based approach, a set of words describing various opinions are manually collected with their known orientations which is a mean of a pre-processing task. The set then grows gradually by searching widely used lexicon dictionary tool such as WordNet or Sentiful[15]. The main objective of the machine learning techniques is in developing the algorithm which optimizes the performance of a system using the training data. This provides a solution of the sentiment classification in two sequential steps:

1)Create and train the model utilizing preparing set information for example effectively marked information[15].

2) Classifying the unlabeled or unclassified data based on the trained or skilled model[16].

Machine learning techniques are further classified into supervised and unsupervised techniques. For carrying out sentiment analysis, typically the supervised machine learning techniques are used as we deal only with the subjective data.AI strategies are additionally characterized into managed and unsupervised techniques. Supervised AI procedures exceptionally rely upon preparing data.Based on the gave preparing information, the classifier will order the rest information for example test data[17]. Countless AI calculations, for example, Logistic

Regression, Naïve Bayes, Decision Tree, Support Vector Machine (SVM), Random Forest, Maximum Entropy, and Bayesian Network are utilized for assessment investigation. Picking a fitting calculation for the information chose is a pivotal advance.

IV. SUPERVISED MACHINE LEARNING ALGORITHMS

Looking from outside, the in-depth study of various supervised machine learning algorithms, it's been observed that many of the following machine learning algorithms are diversely used and gave a average mean of accuracies in most of the domains as well as with many different types of information[18]. They provide constant speed of classification process irrespective of the input size of the data handled.

A. Naïve Bayes (NB) Technique

Naïve Bayes classification was a simple probability based classifier which uses the idea of mixture models for performing the classification[19]. These mixture model depends on the assumption that each of the predefined classes will be one of the mixture components. The various components of these mixture model points the probability nature of presence of any term to the specific component[20]. So, they might be also called as generative classifiers. Naïve Bayes classifier uses the concept of Bayes Theorem in which it finds the maximum aspect of the probability of any word matching to a particular or predefined class.

Given a Hypothesis A and evidence B, Bayes' Theorem states that the relationship between the probability of Hypothesis before getting the evidence P(A) and the probability of the hypothesis after getting the evidence   P (A | B) is :

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

B. Maximum Entropy Method

This entropy method depends on the probability estimation and distribution technique to perform many classifications[20]. Here, in this technique, at the first step, the categorized sets are converted into vectors using any of the encoding techniques available. At the next step, the vector which has been encoded is used to calculate or compute the weights for each of the features which has been extracted that can support collectively to determinethe most prospective label for a feature set. It is being used in various natural language processing works such as text classification. It relies on the probability based approach like Naïve Bayes[21]. The most fundamental concept of maximum entropy is if the most important information about the data is not known, then the distribution must be extremely uniform[21]. This constraint leaves out the probability of non-uniform distribution.

C. Random Forest

This Random Forest classifier is a tree-based classifier model in which it consists of diverse classification trees which can be used in predicting the class label of a data point which are given based on the categorically dependent variable[22]. For a given single data point, each and every tree do votes for a specific class label and that class label gaining the maximum votes will be assigned to that data point[23].The blunder rate of this classifier relies upon the connection or relationship among any two trees in the woodland notwithstanding the quality of unmistakable or individual tree in the backwoods. So as to limit the blunder rate, the trees ought to be solid and the level of associativity ought to be as less as possible[24]. In the classifier tree, the inward hubs are spoken to as the highlights, the edges leaving a

hub are spoken to as tests on the element's weight, and the leaves are spoken to as class classifications. It performs order fundamental from the root hub and moves gradually descending until a leaf hub is distinguished. The report is then arranged in the classification that names the leaf node[24]. This calculation is utilized in numerous utilizations of discourse and language handling.

V. METHODOLOGY

The project had been divided into two phases. First, the literature survey is conducted which is then followed by the system development. The literature study involves conducting various studies on different sentiment analysis methods and techniques that are being currently used[25]. In second phase, the application needs and their functionalities are described prior to the development. Also, architecture and interface design are also identified. In the development of Twitter Sentiment Analysis method, several tools are used, such as Python Shell 2.7.2 and Notepad.

VI. THE PROPOSED SYSTEM

Generally, to build a sentiment analysis model, there exists a three stages methodology which consists of, first building the sentiment words, then extracting the features and then classifying those set of words to predict the positive outcomes[26]. This technique has been employed on two different algorithms on which the better optimized results has been inferred. The proposed paper on twitter data is based on two important things visually the Data Extraction, pre-processing of extracted data and classification[27]. In proposed system we go for machine learning algorithm to analyse whether the given sentence is positive or negative, its accuracy level, and also the classification rate. We use Adaboost and MLP(neural network) ML algorithms.
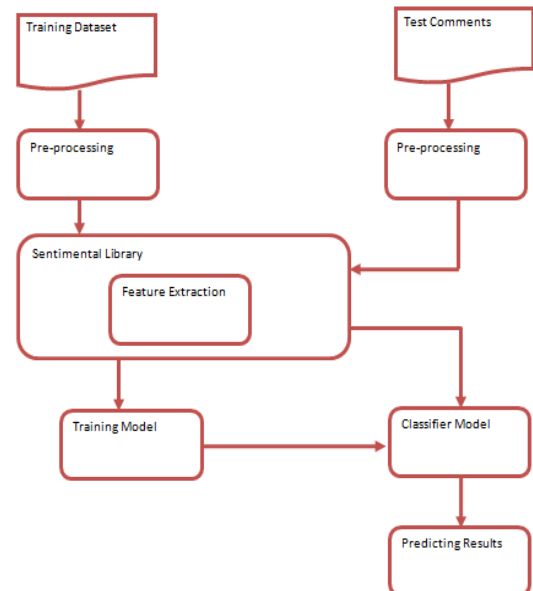
The given input data is downloaded from an open source site named kaggle. The tweets of different election candidates mainly trump and Clinton are given as an input data. Further it is splitted into train data and test data. Training data is nothing but it is used to train and fit the model basically to meet the parameters wherein the test data is used only to assess performance of the model[26]. Training data's output is available to the model whereas testing data is the unseen data for which predictions have to be made. Train data undergoes pre-processing , where the preprocessing use NLP(natural language processing technique), it is the technique used to understand the computer and manipulate the human language[27]. From the NLP, we check if any string is equal or the length of the sentence is equal or not. Here the given tweets contains url id . There is no need for any url that's why they are eliminated. these all things are taken in pre – processing stage.In feature extraction, the tfidf vectorization has been used.Tweets contains different features like symbols, hashtag, punctuation etc. These features are extracted

by using tfidf vectorisation. This converts the all the sentences into vector formats like countable ones[28]. And also it converts a collection of raw documents to a matrix of TF-IDF features. This is equivalent to CountVectorizer followed by TfidfTransformer. The Tokenize and word2vec are imported. They extract the feature for processing things like .counting the word in a statements.,vocabulary etc . The tfidf function is applied on training data set. It takes the part of speech value, Part-Of-Speech tagging (or POS tagging, for short) is one of the main components of almost any NLP analysis. The task of POS-tagging is it simply implies labelling words with their appropriate Part-Of-Speech (Noun, Verb, Adjective, Adverb, Pronoun). These same process is taken by the tested data also. Then the algorithm performance is checked. For machine learning algorithm they have used sentence analysis and feature analysis as well as part-of- speech n-grams[30]. AI is a method for distinguishing designs in information and utilizing them to naturally make expectations or decisions.The two primary strategies for AI are relapse and classification[32]. Here we foreseen given assessments of the distinctive tweets and furthermore by a calculation exhibition.

VII.   APPLICATION AREAS OF SENTIMENT ANALYSIS

Sentiment Analysis has gained popularity in recent years and has been applied in many application areas[34]. It has been used in diversified areas like healthcare, the financial sector, sports, politics, hospitality and tourism, and consumer behavior. Some of the growing and emerging application areas will be reviewed in this section. Sentiment analysis has many applications and benefits to your business and organization[35]. It tends to be utilized to give the business significant bits of knowledge into how individuals feel about an item image or administration. At the point when connected to online life channels, it very well may be utilized to

distinguish spikes in assumption, in this manner enabling them to recognize potential item supporters or internet based life influencers[36]. It very well may be utilized to distinguish when potential negative strings are rising internet in regards to your business, along these lines enabling them to be proactive in managing it the more rapidly. Assessment investigation could likewise be connected to the corporate system, for instance, by applying it to an email server, messages could be observed for their general "tone".Social media are channels of communication with many customers these days, and whenever they're unhappy about something related to a specific company, whether or not it's their fault, they'll call them out on Facebook or Twitter or Instagram[37]. Such mentions will appear in their respective dashboard with a flashing red colour, and they can better start engaging them as soon as they are there.Individuals these days anticipate that brands should react via web-based networking media very quickly, and if the organization isn't sufficiently brisk, they should see them proceeding onward to separate contenders as opposed to sitting tight for their answer[40].

VIII.   CONCLUSION AND DISCUSSION

Sentiment analysis had been proven effectively in predicting the people attitude by analysing many big data's. In this particular research, a novel adaptable approach is presented which aims in extracting the people opinion about a specific concept based on social media contents. The technique that has been proposed contains building a dictionary of words or sentiments by polarity based on a very small group of positive and negative hashtags related to a given subject, and then, classifying those posts into several classes and thereby balancing the sentiment weight factors by using new techniques such as uppercase letters and the repetition of letters in a word. In order to test this model, a case study has been conducted for the 2016 US presidential election to go examine

our model step by step individually to guess which of the candidates was the favorite among people. The performance results had shown a promising outcomes compared to the previous model. However, the proposed approach distinguishes the impact degree of the different metrics in order to accentuate a feeling. Second, a larger Twitter dataset has been used. Third, the system is the prototype developed to assess the ability of dynamic dictionary which is constructed automatically using small samples. As future work, we intent to tackle these three limitations by proposing a more global and efficient model using larger volumes of data.

IX. REFERENCES

[1] R. Addo-tenkorang and P. T. Helo, "Big data applications in operations/supply-chain management: A literature review," Comput. Ind. Eng., vol. 101, pp. 528–543, Nov. 2016

[2] R. Piryani, D. Madhavi, and V. K. Singh, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015," Inf. Process. Manage., vol. 53, no. 1, pp. 122–150, 2017.

[3] A. Pak,and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," Special Issue of International Journal of Computer Application, France:University Paris-Sud, 2010.

[4] "Three Cool and Inexpensive Tools to Track Twitter Hashtags", June 11,2013.[Online].Available [Accessed: 19-Oct-2015].

[5] K. Ghag and K. Shah, "Comparative analysis of the techniques for Sentiment Analysis", in Int. Conf. on Advances in Technology and Engineering, 2013,.

[6] K. Khan, B. Baharudin, A. Khan and F. Malik, "Mining Opinion from Text Documents: A Survey", Digital Ecosystems and Technologies, 2009.

[7] B. Pang, and L. Lee, "Opinion mining and sentiment analysis," 2nd workshop on making sense of Microposts. Ithaca: Cornell University. Vol.2(1), 2008.

[8] E. Kouloumpis, T. Wilson, and J. Moore, "Twitter Sentiment Analysis:The Good the Bad and theOMG!", (Vol.5). International AAAI, 2011.

[9] A. Sarlan, C. Nadam and S. Basri, "Twitter Sentiment Analysis", in Int. Conf. on Information Technology and Multimedia, 2014.

[10] Cui A, Zhang M, Liu Y, Ma S, 2011. Emotion tokens: Bridging the gap among multilingual twitter sentiment analysis. In: Asia Information Retrieval Symposium. Springer

[11] M.Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, " Lexicon Based Methods for Sentiment Analysis," Association for Computational Linguistics, 2011

[12] Tang J, Hu X, Gao H, Liu H, 2013. Exploiting local and global social context for recommendation. In: Ijcai. pp. 2712–2718.

[13] E. Cambria, "Affective computing and sentiment analysis," IEEE Intell. Syst., vol. 31, no. 2, pp. 102–107, Mar./Apr. 2016.

[14] D. H. Wolpert and W. G. Macready, "No free lunch theorems for search," Santa Fe Institute, Santa Fe, NM, USA, Tech. Rep. SFI-TR-05- 010, 1995.

[15] Rosenthal, Sara, NouraFarra, and PreslavNakov. "SemEval-2017 task 4: Sentiment analysis in Twitter." In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), pp. 502-518. 2017.

[16] Medhat, Walaa, Ahmed Hassan, and HodaKorashy. "Sentiment analysis algorithms and applications: A survey." Ain Shams Engineering Journal 5, no. 4 (2014)

[17] Z. Kechaou, B. M. Ammar and A. M. Alimi, "Improving e-learning with sentiment analysis of users' opinions", in Global Engineering Education Conference (EDUCON), 2011

[18] E. H.-J. Kim, Y. K. Jeong, Y. Kim, K. Y. Kang, and M. Song, "Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news," J. Inf. Sci., vol. 42, no. 6, pp. 763–781, 2016.

[19] H. Saif, Y. He and H. Alani, "Alleviating Data Scarcity for Twitter Sentiment Analysis". Association for Computational Linguistics, 2012.

[20] R. Dong, M. P. O'Mahony, M. Schaal, K. McCarthy, and B. Smyth, "Combining similarity and sentiment in opinion mining for product recommendation," J. Intell. Inf. Syst., vol. 46, no. 2, pp. 285–312, 2016

[21] E. D. Avanzo and G. Pilato, "Mining social network users opinions 'to aid buyers' shopping decisions," Comput. Hum. Behav., vol. 51, pp. 1284–1294, Oct. 2015.

[22] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," Commun. ACM, vol. 59, no. 7, pp. 96–104, 2016.

[23] X. Zhang, H. Fuehres, and P. A. Gloor, "Predicting stock market indicators through Twitter 'I hope it is not as bad as I fear,'" Procedia-Social Behav. Sci., vol. 26, pp. 55–62, Jan. 2011.

[24] N. Li and D. D. Wu, "Using text mining and sentiment analysis for online forums hotspot detection and forecast," Decis. Support Syst., vol. 48, no. 2, pp. 354–368, 2010

[25] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R.Passonneau, "Sentiment Analysis of Twitter Data," Annual International Conferences. New York:Columbia University, 2012.

[26] S. Pei, L. Muchnik, J. S. Andrade, Jr., Z. Zheng, and H. A. Makse, "Searching for superspreaders

of information in real-world social media," Sci. Rep., vol. 4, Jul. 2014, Art. no. 5547.

[27] A.C.E.S Lima. and L.N. de Castro, "Automatic sentiment analysis of Twitter messages", in 4 th Int. Conf. on Computational Aspects of Social Networks (CASoN), 2012

[28] Wasserman S, Faust K, 1994. Social network analysis: Methods and applications. Vol. 8. Cambridge university press

[29] V. Sehgal and C. Song, "SOPS: Stock Prediction Using Web Sentiment", in 7th IEEE Int. Conf. on Data Mining Workshop, 2007

[30] P. Nakov, Z. Kozareva, A. Ritter, S. Rosenthal, V. Stoyanov, T. Wilson, Sem Eval-2013 Task2:Sentiment AnalysisinTwitter (Vol.2,pp. 312-320 ,2013.

[31] Tan C, Lee L, Tang J, Jiang L, Zhou M, Li P, 2011. User-level sentiment analysis incorporating social networks. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM.

[32] Thelwall Mike. 2010. Emotion Homophily in Social Network Site Messages. First Monday 15(4)

[33] A. Pak,and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," Special Issue of International Journal of Computer Application, France:Universitede Paris-Sud, 2010.

[34] S. Bahrainian and A. Dangel, "Sentiment Analysis using Sentiment Features", in Int. joint Conf. of Web Intelligence and Intelligent Agent Technologies, 2013.

[35] Mei Q, Ling X, Wondra M, Su H, Zhai C, 2007. Topic sentiment mixture: modeling facets and opinions in weblogs. In: Proceedings of the 16th international conference on World Wide Web. ACM

[36] Pang, Bo, and Lillian Lee. "Opinion mining and sentiment analysis." Foundations and Trends® in Information Retrieval 2, no. 1–2 (2008)

[37] N. Altrabsheh, M. Cocea and S. Fallahkhair, "Sentiment analysis: towards a tool for analysing real-time students feedback", in 26th International Conference on Tools with Artificial Intelligence, 2014.

[38] A. M. Azmi and S. M. Alzanin, "'Aara'—A system for mining the polarity of Saudi public opinion through e-newspaper comments," J. Inf. Sci., vol. 40, no. 3, pp. 398–410, 2014

[39] P. Sobkowicz, M. Kaschesky, and G. Bouchard, "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the Web," Government Inf. Quart., vol. 29, no. 4, pp. 470–479, 2012

[40] B. Gokulakrishnan, P. Plavnathan, R. Thiruchittampalam, A. Perera and N. Prasath, "Opinion Mining and Sentiment Analysis on a Twitter Data Stream", in Int. Conf. on Advances in ICT for Engineering Regions, 2012, pp. 182-188.

[41] C. Chiu, N.-H. Chiu, R.-J. Sung, and P.-Y. Hsieh, "Opinion mining of hotel customer-generated contents in Chinese weblogs," Current Issues Tourism, vol. 18, no. 5, pp. 477–495, 2015.

[42] A. M. Azmi and S. M. Alzanin, "'Aara'—A system for mining the polarity of Saudi public opinion through e-newspaper comments," J. Inf. Sci., vol. 40, no. 3, pp. 398–410, 2014.

[43] Tan C, Lee L, Tang J, Jiang L, Zhou M, Li P, 2011. User-level sentiment analysis incorporating social networks. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 1397–1405

[44] M. del Pilar Salas-Zàrate, E. López-López, R. Valencia-García, N. Aussenac-Gilles, Á. Almela, and G. Alor-Hernández, "A study on LIWC categories for opinion mining in Spanish reviews," J. Inf. Sci., vol. 40, no. 6, pp. 749–760, 2014.

[45] K. Pasupa, P. Netisopakul, and R. Lertsuksakda, "Sentiment analysis of Thai children stories," Artif. Life Robot., vol. 21, no. 3, pp. 357–364, 2016.

Cite this article as :