ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 6, June 2016

Considering Two Sides of One Review for Sentiment Classification

Lincy. W[1], Naveen Kumar. A[2]
[1]PG Scholar,
Department of Computer Science and Engineering,
Saveetha Engineering College,Tamil Nadu, India.
[2]Assistant Professor (SG),
Department of Computer Science and Engineering,
Saveetha Engineering College, Tamil Nadu, India.

***Abstract*—Online reviews available on the internet are majorly used in sentiment analysis and opinion mining to determine the sentiment expressed in the text. Sentiment classification is a basic task in sentiment analysis, with its aim to classify the sentiment of a given text as either positive or negative. Nowadays Bag-of-words (BOW) is mostly used for sentiment analysis in statistical machine learning approaches. Although the BOW model is very simple and quite efficient in topic-based text classification, it is actually not very suitable for sentiment classification. It cannot address the polarity shift problem. The proposed model called Dual Sentiment Analysis (DSA), is address this problem for sentiment classification. Initially, data expansion technique is to be used to create a sentiment-reversed review for each training and test review. On this basis, the Dual Sentiment Analysis (DSA) algorithm is to be applied in two stages. First, dual training algorithm to make use of original and reversed training reviews in pairs for learning a sentiment classifier. Second, a dual prediction algorithm to classify the test reviews by considering two sides of one review. To remove DSA's dependency on an external antonym dictionary for review reversion, a pseudo antonym dictionary is developed based on corpus-based method. At last, the statistical machine learning algorithms (naıve Bayes, maximum entropy classifier, and support vector machines) will be used to train a sentiment classifier.**

Keywords*—Natural language processing, Sentiment analysis, Opinion mining, Machine learning.

I. INTRODUCTION

Online reviews available on the internet are majorly used in sentiment analysis and opinion mining to determine the sentiment expressed in the text [2], [16]. Sentiment analysis is becoming a hotspot in the field of data mining and natural language processing.

Sentiment classification is a basic task in sentiment analysis, with its aim to classify the sentiment (e.g., positive or negative) of a given text [4], [7], [8], [25]. The general practice in sentiment classification follows the techniques in traditional topic-based text classification, where the bag of-words (BOW) model is typically used for text representation. In the BOW model, a review text is represented by a vector of independent words.

Although the BOW model is very simple and quite efficient in topic-based text classification, it is actually not very suitable for sentiment classification. Because, it breaks the syntactic structures, disrupts the word order, and discards some semantic information. One of the most well-known difficulties in BOW is the polarity shift problem [25], [39], [42]. Polarity shift is a kind of linguistic phenomenon which can reverse the sentiment polarity of the text. Negation is the important factor to be considered in polarity shift. By adding a negation word in front of the positive word, the sentiment of the text will be reversed from positive to negative. For example, by adding a negation word "don't" to a positive text "I like this movie" in front of the word "like", it is reversed from positive to negative "I don't like this movie".Several approaches have been proposed in the literature to address the polarity shift problem [16], [17], [21].

II. RELATED WORK

A. Sentiment analysis

According to the levels of granularity, tasks in sentiment analysis can be classified into four categorizations: document level, sentence level, phrase level, and aspect level sentiment analysis [4], [33]. There are many challenges have been figured out in sentiment analysis [4], [7], [23], [31], [35], [43].

Sentiment analysis is not a new topic for the recent years. In the past decade, there has seen a rapid

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 6, June 2016

growth in the field of sentiment analysis from text data. Hatzivassiloglou and McKeown [39] have developed an algorithm for predicting semantic orientation. Their algorithm performs well, but it is implemented for isolated adjectives, rather than phrases containing adjectives or adverbs. Then Point-wise Mutual Information-Information Retrieval algorithm(PMI-IR) is employed to estimate the semantic orientation of a phrase [40]. Peter D. Turney presented a simple unsupervised learning algorithm for classifying reviews as recommended (thumbs up) or not recommended (thumbs down). The classification of a review is calculated by the average semantic orientation of the phrases in the review that contain adjectives or adverbs [21].

In the machine learning methods, the text is represented by Bag-Of-Words (BOW). The Bag Of Words method is typically used for text representation. In BOW text is represented by vector of words. The DS method is proposed by [26], where "NOT" is attached to the words in the scope of negation, e.g., "The Samsung battery life is not good" is converted to "The Samsung battery life is good-NOT". Light-weight Semantic Similarity method is proposed [41] where each text is split up into two parts: polarity-shifted and polarity-unshifted, based on which two component classifiers are trained and combined for sentiment classification. To overcome the polarity shifting problem, Dual Sentiment Analysis method is introduced with selective data expansion [25]. In this algorithm, for each positive/negative review the polarity shifted review is created with the support of Word-Net.

There were still some approaches that addressed polarity shift without complex linguistic analysis and extra annotations. For example, Li and Huang [19] proposed a method first to classify each sentence in a text into a polarity-unshifted part and a polarity-shifted part according to certain rules, then to represent them as two bags-of-words for sentiment classification. Li et al. [21] further proposed a method to separate the shifted and unshifted text based on training a binary detector. Classification models are then trained based on each of the two parts. An ensemble of two component classifiers is used to provide the final polarity of the whole text.

There exists many techniques that explore sentiment analysis which deals with different levels of granularity. In the document-level and sentence-level classification, there are two types of methods used in literature; term counting methods and machine learning methods. In term counting methods, the overall orientation of a text is calculated by summing up the orientation scores of content words in the text, based on manually-collected or external lexical resources [20], [21], [25]. In the machine learning methods, Naïve Bayes Classifier, Support Vector Machine(SVM) and Entropy Classifier are used to train the classifier [9], [22], [25], [34], [42].

III. THE PROPOSED SCHEME

The method called Dual Sentiment Analysis by using Mutual Information(MI) is proposed in this paper.

A. Architectcure Design

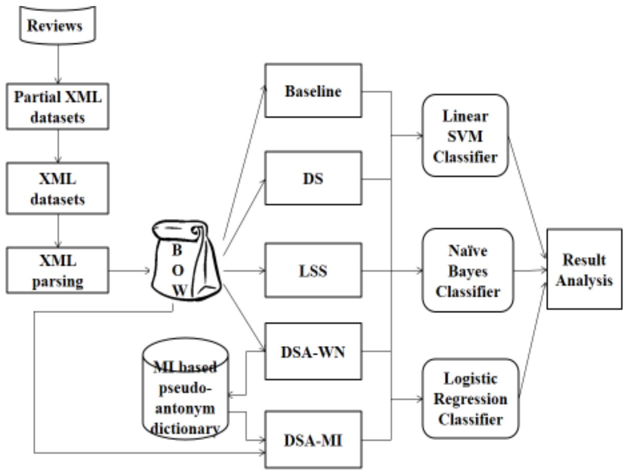

Fig.1 Architecture Diagram

Initially, the collected datasets are in Partial XML format. Then it is converted into XML format. Any user defined tag is added with proper xml tags to convert it into XML format. After the XML parsing, only the REVIEW-TEXT is extracted and saved in separate text document. Each text document contains the unique reviews.By using the method called tokenization, all the words in all the documents are tokenized and saved in a single text document. Then, to find the frequency count of words in each document, HashMap is used.

The following five systems that are proposed in the literature have to be evaluated with the aim at addressing polarity shift.

1) *BOW*:The general practice in sentiment classification follows the techniques in traditional topic-based text classification, where the bag-of-words (BOW) model is typically used for text representation. In the BOW model, a review text is represented by a vector of independent words.
2) *DS:*In this, "NOT" is attached to the words in the scope of negation, e.g., "The book is not interesting" is converted to "The book is interesting-NOT".
3) *LSS:*In LSS, each text is split up into two parts: polarity-shifted and polarity- unshifted, based on which two component classifiers are trained and combined for sentiment classification. To our knowledge, this is the state-of-the-art approach of

considering polarity shift without using external resources.

4) *DSA-WN:*The DSA model with selective data expansion and the WordNet antonym dictionary.
5) *DSA-MI:*The DSA model with selective data expansion and the MI-based pseudo-antonym dictionary.

B. Data Expansion Technique

Given an original sample and an antonym dictionary (e.g., WordNet), a polarity-opposite sample is generated artificially according to the following rules:

1) *Sentiment word reversion:* All sentiment words out of the scope of negation are reversed to their antonyms;
2) *Handling negation*: If there is a negation expression, first detect the scope of negation, and then remove the negation words (e.g., "no", "not", and "don't"). The sentiment words in the scope of negation are not reversed;
3) *Label reversion*: The class label of the labeled sample is also reversed to its opposite (i.e., Positive to Negative, or vice versa) as the class label of newly generated samples (called polarity-opposite samples).

A simple example is shown to explain the generation process.

The original sample
Text: I don't like this movie. Worst story.
Label: Negative

According to Rule 1, "worst" is reversed to its antonym "best"; According to Rule 2, the negation word "don't" is removed, and "like" is not reversed; According to Rule 3, the class label Negative is reversed to Positive. Finally, an artificial polarity-opposite sample is generated:

The generated opposite sample
Text: I like this movie. Best story.
Label: Positive

All samples in the training and test set are reversed to their polarity-opposite versions.

IV. IMPLEMENTATION AND METHODOLOGY

A. Dataset Collections

For polarity classification, the Multi-Domain Sentiment Datasets are used. They contain product reviews taken from Amazon.com including four different domains: Book, DVD, Electronics and Kitchen. Each of the reviews is rated by the customers from Star-1 to Star-5. The reviews with Star-1 and Star-2 are labeled as Negative, and those with Star-4 and Star-5 are labeled as Positive. Each of the four datasets contains 1,000 positive and 1,000 negative reviews.

TABLE 1
DATASET COLLECTION

Dataset	#positive	#negative
Book	1,000	1,000
DVD	1,000	1,000
Electronics	1,000	1,000
Kitchen	1,000	1,000

B. Data Pre-Processing Phase

The data pre-processing phase incorporates steps to modify the data into a form that can be easily and effectively used by the data mining phase [11], [15], [18], [27], [32]. The steps involved are described here.

1) XML parsing

Initially, the datasets are in partial XML format. Then, it is converted into fully XML format by adding XML tags. From this XML datasets, we need the review content alone. So, XML parser is applied to get the reviews.

Fig.2. XML parsing

2) Stop Word Removal

Stop words are the most commonly occurring words in any sentence [12]. These words are filtered prior to automatic text analysis as they do not add any significance during further processing [13]. For removal of stop word, a stop word list is built consisting of 22249 words which include articles, pronouns, proper nouns such as name of the countries, name of the personalities etc. If the sentence contains words that are present in the stop word list, then they are eliminated. Example of sentences revealing positive and negative emotions are shown in fig.2

Emotion	ID	Example
Positive	1	Good Acting. Nice story. Climax was unpredictable. Everyone has done their work nicely. Overall film was awesome
Negative	2	I don't like this book. It is boring. Last chapter is completely boring.

Fig. 2. Example of sentences expressing positive and negative emotions

Illustration of the process of stop word removal is provided in Fig. 3.

ID	Sentence after stop word removal
1	Good Acting Nice unpredictable done nicely awesome
2	don't like boring boring

Fig. 3. Illustration of stop word removal

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 6, June 2016

The output of this process is given as input to stemming.

4) Stemming

English words can be inflected which means that they can change their form to express differences in number, tense, gender, person, case, aspect, mood etc. In text analysis and natural language processing, these words are reduced to their root or base form to improve further processing. This process is termed as stemming [14]. Stemming is performed on the output obtained from the stopword removal process and the result is portrayed in fig. 4.

ID	Result after Stemming
1	Good Act Nice Unpredictable Do Nice Awesome
2	Don't Like Bore Bore

Fig. 4. Result of Stemming

These identified root word forms the output of the data pre-processing phase. They are provided as input to the data mining phase, steps of which are explained in the following sub-section.

C. Dual Sentiment Analysis

DSA algorithm contains two main stages called Dual Training and Dual Prediction.

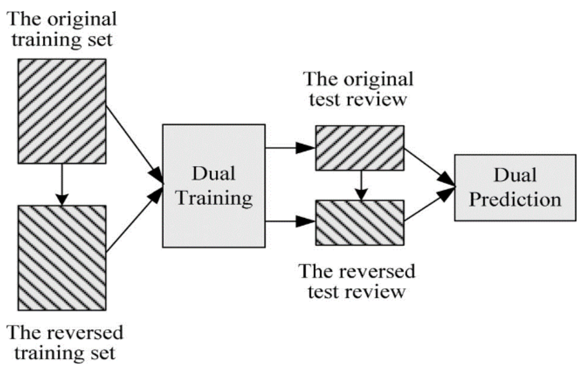

Fig. 2. The process of dual sentiment analysis. The rectangle filled with slash denotes the original data, and the rectangle filled with backslash denotes the reversed data.

1) *Dual Training:*In the training stage, all of the original training samples are reversed to their opposites. It is referred as "original training set" and "reversed training set" respectively. In data expansion technique, there is a one-to-one correspondence between the original and reversed reviews. The classifier is trained by maximizing a combination of the likelihoods of the original and reversed training samples. This process is called dual training.
2) *Dual Prediction:*In the prediction stage, for each test sample x, a reversed test sample ~x is created. The aim is not to predict the class of ~x. But instead, ~x is used to assist the prediction of x. This process is called dual prediction.

In DP, predictions are made by considering two sides of one review:

- When we want to measure how positive a test review x is, we not only consider how positive the original test review is, but also consider how negative the reversed test review is;
- Conversely, when we measure how negative a test review x is, we consider the probability of x being negative as well as the probability of ~x being positive.

To reduce DSA's dependency on an external antonym dictionary, a corpus-based method is to be developed for constructing a pseudo-antonym dictionary. The pseudo antonym dictionary is language-independent and domain adaptive. It makes the DSA model possible to be applied into a wide range of applications.

V. EXPERIMENTS AND RESULTS

The performance of the classifiers was evaluated through various evaluation metrics. Ten fold cross validation was performed for assessment. Metrics such as accuracy, precision, recall and f-measure were computed [3]. Accuracy signifies the number of correctly predicted instances to the total number of instances [1].

$$Accuracy = \frac{Number\ of\ correct\ results}{Total\ number\ of\ instances}$$

Precision refers to the fraction of number of correct results to the number of all returned results.

$$Precision = \frac{Number\ of\ predicted\ correct\ results}{Number\ of\ predicted\ results}$$

Recall is the number of correct results divided by the number of results that should have been returned.

$$Recall = \frac{Number\ of\ predicted\ correct\ results}{Number\ of\ actual\ correct\ results}$$

F-measure denotes the harmonic mean of precision and recall. Mathematically, it is represented as follows.

$$FMeasure = \frac{2 * Precision * Recall}{Precision + Recall}$$

The classification rules evolved by the best performing classifier are used for prediction of emotions from reviews [2]. The experiments and the associated results are discussed in the following section.

Sentiment classification from reviews available on the internet have been attempted in this work. Various experiments were conducted in view to attain maximum possible accuracy. The impact of feature selection and ensemble techniques were analyzed. Initially various classifiers namely Random Tree (RT), Reduced Error Pruning Tree (REPT), C4.5, Grafted C4.5, FT and Classification and Regression Trees (CART) were employed to classify the emotions in the news articles. Then, feature selection was adopted and classification was attempted on the reduced set of features. Table 2 presents the accuracy results obtained with and without feature selection.

TABLE 2
RESULTS OF ACCURACY PREDICTIONS WITH AND WITHOUT FEATURE SELECTION

Classifier	Accuracy	
	Without Feature Selection	With Feature Selection
RT	76.37	81.67
C4.5	86.08	86.83
Grafted C4.5	86.46	86.04
REPT	85.02	86.87
FT	85.88	86.25
CART	85.81	86.73

V. CONCLUSION

In this paper, DSA algorithm is proposed to address the polarity shift problem in sentiment classification.In data expansion technique, reversed reviews are created. There are two stages in DSA algorithm: Dual Training and Dual Prediction. In the training stage, all of the original training samples are reversed to their opposites. In the prediction stage, for each test sample, a reversed test sample is created. Finally, the statistical machine learning algorithms (such as naıve Bayes, maximum entropy classifier, and support vector machines) used to train a sentiment classifier. Based on the DSA-WN model, a corpus-based method for constructing a pseudo-antonym dictionary is created. In future, the antonym dictionary can be developed manually with more accurate words and phrases.

REFERENCES

[1] A.McCallum and K. Nigam, “A comparison of event models for naive Bayes text classification,” in Proc. AAAI Workshop Learn. Text Categorization, 1998, pp. 307–314.
[2] Alexander Pak, Patrick Paroubek, “Twitter as a Corpus for Sentiment Analysis and Opinion Mining”, 2010.
[3] Amandeep Kaur, Vishal Gupta, “A Survey on Sentiment Analysis and Opinion Mining Techniques”, Journal Of Emerging Technologies In Web Intelligence, Vol. 5, No. 4, November 2013.
[4] Bing Liu, Lei Zhang, “A Survey Of Opinion Mining And Sentiment Analysis”, Book.
[5] Despo Georgiou, Andrew MacFarlane, “Extracting Sentiment from Healthcare Survey Data: An Evaluation of Sentiment Analysis Tools”, Science and Information Conference 2015 July 28-30, 2015.
[6] Emma Haddia, XiaohuiLiua, Yong Shib, “The Role of Text Pre-processing in Sentiment Analysis”, Information Technology and Quantitative Management (ITQM2013), Procedia Computer Science 17 (2013) 26 – 32.
[7] G.Vinodhini, RM.Chandrasekaran, “Sentiment Analysis and Opinion Mining: A Survey “,International Journal of Advanced Research in Computer Science and Software Engineering”, Volume 2, Issue 6, June 2012.
[8] GOA project Biograph from the University of Antwerp Cluster of Excellence, MMCI, Saarland University, “Negation and Speculation in Natural Language Processing”, NeSp-NLP2010.
[9] Hailong Zhang, WenyanGan, Bo Jiang, “Machine Learning and Lexicon based Methods for Sentiment Classification: A Survey”, 11th Web Information System and Application Conference, 2014.
[10] Harshali P. Patil, Mohammad Atique, “Sentiment Analysis for Social Media: A Survey”, IEEE, 2015.
[11] Hatzivassiloglou, V., & McKeown, K.R. 1997. “Predicting the semantic orientation of adjectives.” Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL (pp. 174-181). New Brunswick, NJ: ACL.
[12] ImeneGuellil, Kamel B, “Social Big Data Mining : A Survey Focused on Opinion Mining and Sentiments Analysis”, IEEE, 2015.
[13] Khaled Ahmed, Neamat El Tazi, Ahmad Hany Hossny, “Sentiment Analysis Over Social Networks: An Overview”, IEEE International Conference on Systems, Man, and Cybernetics, 2015.
[14] Kim Schouten and Flavius Frasincar, “Survey on Aspect-Level Sentiment Analysis”, Ieee Transactions On Knowledge And Data Engineering, Vol. 28, No. 3, March 2016.
[15] Mark E. Larsen, Tjeerd W. Boonstra, Philip J. Batterham, Bridianne O’Dea, Cecile Paris, and Helen Christensen, “We Feel: Mapping Emotion on Twitter”, Ieee Journal Of Biomedical And Health Informatics, Vol. 19, No. 4, July 2015.
[16] Mohammad SadeghHajmohammadi, Roliana Ibrahim, Zulaiha Othman, “Opinion Mining and Sentiment Analysis: A Survey”, International Journal of Computers & Technology Volume 2 No. 3, June, 2012.
[17] Ms. KrantiVithalGhag, Dr.Ketan Shah, “Comparative Analysis of Effect of Stopwords Removal on Sentiment Classification”, IEEE International Conference on Computer, Communication and Control (IC4-2015).
[18] P. Gonçalves and M. Araújo, “Comparing and combining sentiment analysis methods,” In Proceedings of the first ACM conference on Online social networks. ACM., pp. 27-38, 2013.
[19] P. Menaka, A. Prathimadevi, “A Survey on Web Mining and Its Techniques”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 9, September 2015.
[20] P. Turney and M. L. Littman, “Measuring praise and criticism: Inference of semantic orientation from association,” ACM Trans. Inf. Syst., vol. 21, no. 4, pp. 315–346, 2003.
[21] P. Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” Proc. Annu. Meeting Assoc. Comput. Linguistics, 2002, pp. 417–424.
[22] Pang B., L. Lee, and S. Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. In Proceedings of EMNLP-02.
[23] Raisa Varghese, Jayasree M, “A Survey On Sentiment Analysis And Opinion Mining”, IJRET: International Journal of Research in Engineering and Technology, Volume: 02 Issue: 11 | Nov-2013,
[24] Raymond Kosala, H Celestijnenlaan, “Web Mining Research: A Survey”, SIGKKD Explorations, Copyright ACM SIGKKD, July , 2000.
[25] Rui Xia, Feng Xu, ChengqingZong, Qianmu Li, Yong Qi, and Tao Li, “ Dual Sentiment Analysis: Considering Two Sides of One Review”, Ieee Transactions On Knowledge And Data Engineering, Vol. 27, No. 8, August 2015.
[26] S. Das and M. Chen, “Yahoo! for Amazon: Extracting market sentiment from stock message boards,” in Proc. Asia Pacific Finance Assoc. Annu.Conf., 2001.
[27] S. Li, S. Lee, Y. Chen, C. Huang and G. Zhou, “Sentiment classification and polarity shifting,” in Proc. Int. Conf. Comput. Linguistics, 2010, pp. 635–643.
[28] Shoushan Li, Sophia Yat Mei Lee, Ying Chen, Chu-Ren Huang, Guodong Zhou, “Sentiment Classification and Polarity Shifting “,Proceedings of the 23rd International Conference on

Computational Linguistics (Coling 2010), pages 635–643, Beijing, August 2010.

[29] SinemGüven, Mathias Steiner, Niyu Ge, Amit Paradkar, "Understanding the Role of Sentiment Analysis in Contract Risk Classification", IBM T. J. Watson Research Center Yorktown Heights, NY 10598, 978-1-4799-0913-1/14/$31.00 ©2014 IEEE.

[30] Solanki Yogesh Ganeshbhai, Bhumika K. Shah, "Feature Based Opinion Mining: A Survey", IEEE InternationalAdvanceComputing Conference (IACC), 2015.

[31] Subhabrata Mukherjee, "Sentiment Analysis : A Literature Survey ",Indian Institute of Technology, Bombay. Department of Computer Science and Engineering, June 29,2012.

[32] Surbhi Bhatia, Manisha Sharma, Komal Kumar Bhatia, "Strategies forMining Opinions: A Survey", IEEE, 2015.

[33] Theresa Wilson, Janyce Wiebe, Paul Hoffmann, "Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis", Association for Computational Linguistics, 2009.

[34] Turney, P.D. 2001. "Mining the Web for synonyms: PMI-IR versus LSA on TOEFL". Proceedings of the Twelfth European Conference on Machine Learning (pp. 491-502). Berlin: Springer-Verlag.

[35] Vinita Sharma, "Literature Survey", June 29, 2014.

[36] W Medhat, A Hassan, H Korashy, "Sentiment Analysis Algorithm and its Application: A Survey", Ain Shams Engineering Journal, 2014 – Elsevier.

[37] Wiltrud KESSLER Hinrich SCH ¨UTZE, "Classification of Inconsistent Sentiment Words Using Syntactic Constructions", Proceedings of COLING 2012: Posters, pages 569–578, COLING 2012, Mumbai, December 2012.

[38] Y. Bae and H. Lee, "Sentiment analysis of Twitter audiences: Measuring the positive or negative influence of popular twitterers," Journal of the American Society for Information Science and Technology, vol. 63, no. 12, pp. 2521-2535, 2012.

[39] Y. Rao, J. Lei, L. Wenyin, Q. Li, and M. Chen, "Building emotional dictionary for sentiment analysis of online news," World Wide Web, vol. 17, pp. 723-742, Jun. 2014.

[40] Zhaoxia WANG, Victor JooChuan TONG, David CHAN, "Issues of social data analytics with a new method for sentiment analysis of social media data", IEEE 6th International Conference on Cloud Computing Technology and Science, 2014.

[41] Zhaoxia WANG, Victor JooChuan TONG, Xin XIN, HoongChor CHIN, "Anomaly Detection through Enhanced Sentiment Analysis on Social Media Data", IEEE 6th International Conference on Cloud Computing Technology and Science, 2014.

[42] ZohrehMadhoushi, AR Hamdon, S Zainudin, "Sentiment Analysis Techniques in Recent Works", Science and Information Conference 2015 July 28-30, 2015.

[43] Lincy. W, Naveen Kumar. M, "A survey on challenges in sentiment analysis",International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 21 Issue 3 – APRIL 2016.