An Improvement of Sentiment Classification in Cloud Computing using Lexicon Based Dictionary Approach

Dr R Suguna, Assistant Professor, Theivanai Ammal College for Women (Autonomous), Villupuram, India, sugunarajasekar@gmail.com

Dr S Manimekalai, Assistant Professor, Theivanai Ammal College for Women (Autonomous), Villupuram, India, mamekaroshni@gmail.com

Dr S Arulselvarani, Assistant Professor, STET Women's College, Mannargudi, India, arul2catchme@gmail.com

Abstract- Sentiment Analysis is one method of describing the writer's emotional state depending on their written text over the blogs by regarding the polarity of keywords utilized in the writing process. It is one task of gathering entire polarity that may be in the form of positive, negative or neutral towards a subject. However, present classification methods cannot attain text classification with efficient and most of the research achievement may be about a single language. In this paper, it propose a Principle Component Analysis (PCA) in cloud computing using lexicon based dictionary to classifying the user opinions about products on blogs in texts of length of multiple sentences extracted from the online product reviews. It refers to a classification difficulty where the major focus is to predict the polarity of words and after that classify them into positive or negative or neutral sentiment. In this process, the preprocessing method can be applied to remove the unrelated information from the dataset before feature extraction. This proposed approach uses the PCA to extract the features from the product review dataset. The extracted features are matched using lexicon based dictionary method. The SVM classifier is applied to classify the sentiment as positive, negative or neutral with help of lexicon based dictionary. Experimental results of proposed system show that the sentiment analysis method for online product using PCA with SVM classifier method is able to classify the opinions of users accurately.

Keywords* — *Cloud Computing, Feature Extraction, Lexicon Based Dictionary, Principle Component Analysis, Sentiment Analysis, Sentiment Classification, Support Vector Machine.

I. INTRODUCTION

The process of online shopping is one of the part of e-commerce can refers to the process of selling a products or giving a services to online users through the internet. Conversely, it is not an easy issue to estimate those online user's feedbacks about products. While analyzing this rapid growth of online reviews, it can become complicated to classify whether opinion of users about products and service can be satisfied or dissatisfied [1][2]. In addition, as part of enhancing their quality of products or services, organizations such as online shop require to categorize what product features and service that user most incline of. Furthermore, every online review of users about products are not only talk over one topic issue, it can discuss about different characteristics with various range of sentiment such as positive, negative or neutral. To solve this problem sentiment analysis is used, by knowing the emotion and given online feedback's circumstance. Sentiment analysis is considered as the utilization of processing of natural language, computational linguistics and text analysis can be used to recognize and extract sentiment information or data from the source materials [4][5]. Usually, goal of sentiment analysis is to discover the writer's feelings with respect to several relevant topics or the entire related document polarity [6]. Generally, sentiment can falls into the classifications of attitudes, feelings and opinions, and it is subjective that is not depending on the facts. Binary oppositions also can be consideration of as Sentiment such as positive or negative, good or bad and like or dislike. SA is utilized for different purposes such as defining if reviews of product can be positive or negative, if attitudes of bloggers modified as a selection or even finding a suitable content for a position of advertisement. By first using some data, SA can be characteristically executed for feature extraction. With a classification algorithm, these features can be employed together to examine the data stream to establish the particular topic's sentiment [7].

 DOI : 10.18231/2454-9150.2018.1413

Emotions are extensively utilized to express the online users' emotion. It is hard to categorize the emotions at the back the emoticons because the structure in emotions may be uneven. Process of classification can be used to classify instances into their particular classes. Classification can comprise the variables with recognized values to forecast the unidentified or future values of other variables [7][8]. Furthermore, the similar word contains various meanings based on the context. Hence, it is needed to construct sentiment-based domain dictionaries. The dictionary based approach can utilizes an online dictionary to gather features' thesauruses called as seed words to increase set of feature so that unnerved tweets are also employed for classification. In this paper, we propose a SVM with PCA classifier to extract the feature and to classify the reviews from the testing database. In this process, the initial step is calculate exhaustive and contains the preprocessing and classification of an online product reviews of user as the next step can take some time to get an ultimate choice on the sentiment of topic. The training text dataset can be preprocessed by using various filtering methods, with the goal of computing the combinations of several classification algorithms and filtering methods for the assignment of sentiment analysis [10]. In this paper, we introduce an accurate sentiment classifier by State Vector Machine classifier with lexicon based dictionaries for discovering the online product user's review sentiment. A novel approach to text sentiment classification utilizing contextual information or data about reviews specially attained from paragraphs. We mainly focus our discussion on three topics: Feature Extraction of Sentiment, Lexicon based Dictionaries and Sentiment Classification.

II. RELATED WORKS

Jinyan Li, et.al [1] estimated the several popular classification algorithms along with three filtering techniques. Original dataset was shrunk by these three filtering techniques shrunk increasingly, with respect to the frequent terms of a document and related polarity. Coined hierarchical classification was utilized to classify the sentiments based on the emotions. The effects of this approach in different combinations of classification algorithms and filtering techniques were discussed over three sets of contentious online news articles where binary and multi-class classifications is exploited. In nature, the language is very highly agglutinative, Kishorjit Nongmeikapam, et.al [2] has deal about the Manipuri article's sentiment analysis. The document files can be the letters to the editor of a small number of local every day newspapers. The text has been processed by utilizing Conditional Random Field (CRF) method for Part of Speech (POS) tagging. The verbs' lexicon was altered with the sentiment polarity such as Positive or Negative or Neutral manually. With the POS tagger the verbs of each sentence have been discovered and the altered lexicon of verbs can be utilized to inform the sentiment's polarity in the sentence. Xing Mou and Yajun Du [3] presented a novel sentiment classification algorithm for reviews about movie that which integrates multiple lexicons and How also Net sentiment lexicon, Chinese sentiment polarity lexicon (NTUSD) are included, and the extended emotion lexicon and movie filed emotional lexicon, and integrated with the linguistic features to categorize the micro-blog movie reviews' polarity. Haiping Zhang, et.al [4] presented a hybrid technique which have united association rules and point-wise mutual information to extract the product features and after that obtained sentiment dictionary benefit - HowNet was analyzed the opinion orientation expressed on the online shopping's product features.

Dan Li, Jiang Qian [7] presented a Memory (LSTM) which obtains entire sequence information capably. LSTM method has been better compared with the customary RNN language model, in knowing sentiment of long sentences and as a language model, to achieve multi-classification for text emotional attributes, LSTM was utilized. Therefore, although training various emotion models, emotions of sentence belongs to by utilizing these emotion methods. Numerical experiments have shown that it can produce better correctness rate and recall rate than the conservative RNN. Tejasvini Patil and Sachin Patil [6] proposed a novel approach for evaluation of emotion from the text entered by user on social networking sites. An affective words and sentence context analysis methods have been utilized for emotion identification. In this process, a visual image generation approach that makes images according to emotion for assisting the users to effectively state their emotion in text.

III. PROPOSED METHODOLOGY

OVERALL PROCESS OF PROPOSED METHOD

This paper proposes three key steps for sentiment analysis about online product reviews. Figure 1 shows overall process of proposed method for sentiment classification.

Step 1: First step is preprocessing a data, in this process, unwanted noisy data and inconsistent data can be removed from the sample product reviews data. This step removes the URLs, unnecessary questions, copied review commands and special characters to give the high quality reviews about online products.

Step 2: In next step, feature can be extracted from the preprocessed product reviews data. The Principle Component Analysis method is utilized to extract the features of sample input data.

Step 3: Finally, the product reviews are classified using SVM method with lexicon based dictionaries based on the user's thought about products. The sentiments are classified using word matching depending on the lexicon. After that,

 DOI : 10.18231/2454-9150.2018.1413

ISSN : 2454-9150 Vol-04, Issue-10, Jan 2019

the sentiments can be classified as positive, negative, or neutral reviews. The proposed framework has following three steps

- Preprocessing
- Feature extraction using PCA
- Sentiment classification using SVM with lexicon based dictionary

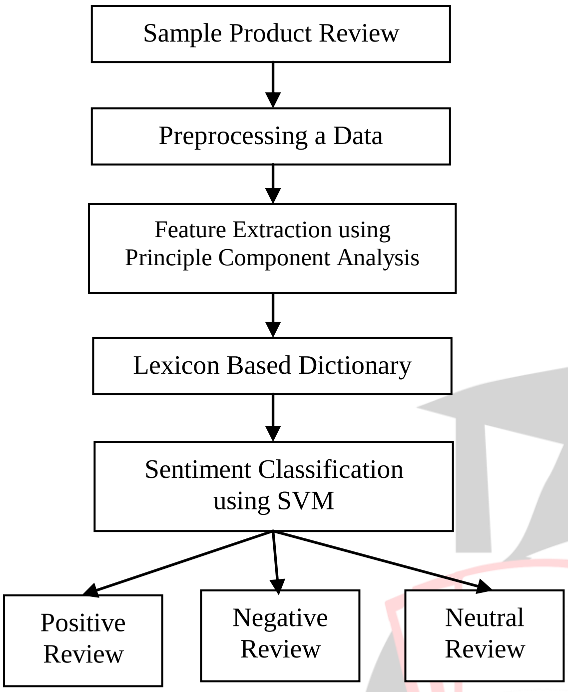

Figure 1 Overall process of proposed method

This paper proposes three key steps for sentiment analysis about online product reviews. Figure 1 shows overall process of proposed method for sentiment classification.

Step 1: First step is preprocessing a data, in this process, unwanted noisy data and inconsistent data can be removed from the sample product reviews data. This step removes the URLs, unnecessary questions, copied review commands and special characters to give the high quality reviews about online products.

Step 2: In next step, feature can be extracted from the preprocessed product reviews data. The Principle Component Analysis method is utilized to extract the features of sample input data.

Step 3: Finally, the product reviews are classified using SVM method with lexicon based dictionaries based on the user's thought about products. The sentiments are classified using word matching depending on the lexicon. After that, the sentiments can be classified as positive, negative, or neutral reviews. The proposed framework has following three steps

- Preprocessing
- Feature extraction using PCA
- Sentiment classification using SVM with lexicon based dictionary

1. Preprocessing

In the preprocessing, to split the each sentence into words or token, it is known as tokenization. Dictionary is constructed to replace the token or word that associated to sentiments and online aspects of online distro. Because some online review commands contains grammatical error and misspell, so this will be critical. To remove the token that not gives any meaning or thought related to the sentiments and online distro aspects. Then it perform the removing URLs, questions, special characters, copied contents from the sample product review database, and filtering process and high quality reviews are also be executed in this paper. Basically, URLs do not give to evaluate the sentiment in the form of text about product review of users. For example, take the sentence from the database "I have closed my user account from, because I received poor product from that website" in fact the above sentence give negative thought. In order to remove this kind of failures to utilize a method to remove URLs from product review database. Next to eliminate the question words from review sentence like what, when, where, which, how etc., to decrease the complexity because these type of questions cannot provide to polarity of sentence.

To remove the Special characters {}, :, ;, [], (), /' in order to remove inconsistency during the polarity task. For instance, "it's good look:" when special characters are not removed, the special characters can be focused with the words and create those words are not available in the dictionary. Therefore, we remove special characters from the sentence to reduce the complexity of dictionary searching. Before filtering the process, they have to convert every uppercase into lowercase letters in the input training dataset. In filtering process, the repeated letters are filtered. Commonly, users may be use repeated letters to expose their feelings about products. For example, "I'm Sooooo Happyyyy about this product"; this sentence gives user's intension of expression. In the sentiwordnet, these words are not presented. Therefore, extra letters of word should be removed.

This kind of removal can follow the rules, a letter should be not repeated more than the three times, and therefore we need to eliminate extra and unwanted letters from the word. We eliminate the copied contents to decrease the memory storage. Sometimes, product users copying another user's commands and paste the commands. This generally occurs in many online product reviews. Finally, unwanted or inconsistent data has been removed by using the preprocessing method before execution of feature extraction.

2. Feature Extraction using PCA

In this paper, we execute the feature extraction using PCA to improve the dimensionality reduction. Dimensionality

 DOI : 10.18231/2454-9150.2018.1413

reduction is used to reduce the computational complexity. Therefore, this dimensionality reduction is critical step to construct a sentiment analysis method in order to eliminate unnecessary and irrelevant comments. This feature extraction method can improve the learning performance, reducing the computational complexity, constructing best models and can give the efficient memory storage. In this paper, feature extraction uses the PCA method and selection of words from sample dataset of online product review. Therefore, the dimensionality reduction can be done by feature extraction using Principle Component Analysis method.

The PCA method is utilized as a dimensionality reduction method for issues of sentiment classification about online product reviews. PCA method can be employed as a tool in analysis of exploratory data and for creating analytical models. By using an eigen-value decomposition of a data covariance matrix or singular value decomposition of a data matrix, the PCA process is performed, generally after mean centering data matrix for every attribute in dataset. Usually, the PCA results can be discussed in terms of scores of component, this type of scores sometimes called a factor scores. In this process, we transform the variable values to the particular data point. In this paper, we use the PCA to transform the multidimensional data into lower dimensional data in online product review dataset. In modern data analysis, PCA can be a standard tool to assumes every variability in a process is utilized in the analysis so it becomes hard to differentiate the significant variable from the less important a data set A_i, (i =1,...,n) can be précised as a linear combination of orthonormal vectors are known as principal components.

$$F(a,x) = y + (aX)\, X^{T} \qquad (1)$$

In above equation, f (a, X) is a valued function of vector, **y** is the mean of the data {w_i}, and X is d × m matrix with orthonormal columns process. The mapping can be written as bi = w_iX can gives a low vectors' dimensional projection w_i if $m < d$

In the maximum variance direction, the first principal component can be an axis. The principal components contains following best properties in the linear functions class f (a, X). The principal components W can give a linear estimation that represents the maximum variance of the original data in a low-dimensional data. They can also be offer the optimal low-dimensional linear representation in the wisdom that the total sum of squared distances from data points to their projections in the space. If the mapping functions I and J can be limited to the linear functions class, the composition I (J(x)) can give the PCA replaces the original variables of an online product review data set with a smaller number of uncorrelated variables. When original data set of dimension D have greatly correlated variables, after that there is an efficient dimensionality, d < D, that describes most of the data. A small number of components of d create it simple to make every dimension with an instinctive meaning. In addition, it is more competent to work on fewer variables in following analysis. Finally, the features are extracted from the online product user's review data. Here, the multidimensionality complexity has been reduced. In this paper, feature is extracted efficiently using PCA method before sentiment classification to easily classify the sentiment analysis.

3. Sentiment classification using SVM with Lexicon based dictionary

After the feature extraction, we use the SVM classifier to classify the sentiments. In our proposed system, the sentiments are classified using SVM with lexicon based dictionary. In general, classification can be used to classify the instances into their particular classes. This process can comprise the variables with recognized values to predict the future values of other variables or unidentified variables in the dataset. In sentiment analysis, we classify the text based on the polarity. Basically, the text classification is to categorize data into predefined classes. In this process, the classes may be in the form of positive and negative, or neutral. The text classification is supervised learning issue. It has been identified that word stop works well as representation unit in recovery of information. This can direct to text's attributed value representation. Every word corresponds to feature with number of times word happens in the review comments, as its value. Therefore, overcome this problem, we use the SVM method with help of lexicon based dictionary to classify the text. This SVM can be universal learner.

In this paper, lexicon based dictionary is used for word matching to classify the sentiments using SVM method efficiently. SVM's significant property is that their capacity to learn is feature space's dimensionality independent. SVM method for sentiment classification for online product review dataset document in original form is not appropriate for learning process. Therefore the data can be converted into format which can matches into input of machine learning algorithm. After that, the process of transformation can be done. Every word can correspond to one dimension and equal words to similar dimension. Here we use dictionary based method by employing help of online dictionary that has antonyms and synonyms for every word. This method utilizes the lexicon-based dictionary which has set of well-lexicons such as WordNet dictionary process in English language. WordNet dictionary can maintains the set of lexical dataset for online product review words and also maintain the sentiment relationship of each words in dataset.

This kind of dictionary based approach utilizes an online dictionary to collect lexicons of features recognized as seed

 DOI : 10.18231/2454-9150.2018.1413

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-04, Issue-10, Jan 2019

words to increase feature set so that frightened text for review of online product can also be exploited for sentiment classification. Here, every word in the lexicon contains positive, negative or neutral polarity that is utilized to classify the sentiments about online product positive, negative or neutral. SVM classify the sentiments using Lexicon-based sentiment dictionary employs lexicon to classify the text by using word matching. The sentiment lexicon may in the form of set of words that contains positive, negative or neutral sentiment and it can be utilized for selection of seed and sentiment categorization as a feature set. The performance of SVM classifier is enhanced as the lexicon has more words that contain strong sentiment to classify the sentence or words. In this method, every tokenized word contains two attributes which can be the base form of word known as lemma and POS tag such as noun(NN), verb(VB), adjective(JJ) and so on. Every tokenized word also stored into the sentiment lexicon with accumulates frequency for sentiment classification. This accumulated frequency can be a term frequency that is the number of presences of a word in the training dataset of online product review.

SVM is a one kind of supervised methods, where in sentiment analysis contains training a sentiment classifier through the frequency of different words appearing in an online product review. Here, Feature vectors can be made for the sentiment classification utilizing SVMs can be constructed to have an important segment of data about the nature of the online product review, we calculate the number of frequent features in the vectors of feature for each pair of dataset with help of lexicon based dictionary. Depending on the frequency of word in the sentiment lexicon, the seed words can be chosen by distinguish and Polarity. Distinguish and Polarity can be estimated as follows:

$$Dis = |\,Positive\ frequency - Negative\ frequency\,) - (Neutral\ frequency) \quad (2)$$

$$Polarity = \frac{Dis}{Positive\ frequency + Negative\ frequency + Neutral\ frequency} \quad (3)$$

Here, Positive frequency can be a term of frequency in positive thought about online product; Negative frequency can be a term of frequency in negative opinion about product and Neutral frequency is a term of frequency in negative and positive opinion about the online product. Equation (2) denotes the gap between positive frequency, negative frequency and neutral frequency. When the distinction between words is increased, then the word contains high term frequency of in the form of positive, negative or neutral sentiments about online product. Equation (3) denotes the word's polarity. When polarity is enhanced, then the word contains stronger sentiment of product. The words cover a definite threshold value can be chosen as the seed words. Finally, the dataset of online product review is classified into positive, negative or neutral using SVM with help of the Lexicon based dictionary.

4. Performance Evaluation

The performance metrics can be used to estimate the results of sentiment classification about online product review are Precision, Recall and F-measure. These types of metrics can be evaluated depending on the values of correctly classified sentence based on the online product user's feelings, true and false positive, true and false negative or true and false neutral emotion.

Precision

The accuracy is the values of correctly classified sentence out of classified sentence based on the user's feelings and the accuracy can be calculated as

$$Precision = \frac{Values\ of\ correctly\ classified\ sentence\ based\ on\ the\ user's feelings}{Values\ of\ all\ classified\ sentence\ based\ on\ the\ user's feelings} \quad ---- \quad (4)$$

Recall

Recall is the values of corrected classified sentence out of the actual whole documents and it is calculated as

$$Recall = \frac{Values\ of\ correctly\ classified\ sentence\ based\ on\ the\ user's\ feeling}{Values\ of\ all\ sentence\ based\ on\ the\ user's\ feelings} \quad ----(5)$$

F1-Measure

Finally, it calculate the F1-measure with precision and recall, which is generally utilized in the retrieval and classification of user sentiments about online product review

$$F1 - Measure = \frac{2(precision\ X\ recall)}{precision + recall} \quad ---(6)$$

Runtime of Classification

The top classifiers attained for every algorithm have been tested against each other utilizing the test online product review dataset to establish the time taken to classify sentiments about online product. As we estimating the time length it takes to process of classify the sentiments in a real-time system, the testing of classification time contained defining the time taken to preprocess the sentiments, extract the features and classify the sentiments. Once the preprocessing, feature extraction and classification steps are finished, the total time taken can be divided by the number of sentiment words in order to get the average time taken to classify sentiment. Table 1 shows the test results of text,

 DOI : 10.18231/2454-9150.2018.1413

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-04, Issue-10, Jan 2019

here, the precision, recall and F1- measure has been tested based on the user's emotions about online product

Table 1 Test results of text

	Case	User's feelings about product	Precision	Recall	F1-Measure
Online Product	Case 1	Positive	0.7801	0.8951	0.8337
		Negative	0.8421	0.7364	0.7857
		Neutral	0.8125	0.7895	0.8008
	Case 2	Positive	0.8421	0.7708	0.8048
		Negative	0.7995	0.8461	0.8221
		Neutral	0.9236	0.8248	0.8714
	Case 3	Positive	0.7509	0.8507	0.7677
		Negative	0.8624	0.7500	0.8022
		Neutral	0.8712	0.8729	0.8720

IV. RESULTS AND DISCUSSIONS

In this section, we give the different experimental results to estimate the performance of existing and proposed method. The data set is collected from mdb online data repository. To run our proposed method based on different matrices and compared with the existing methods. First test the Accuracy, Recall and F1-measure for each method and the compare the runtime to classify the sentiments.

Accuracy

The accuracy of classifier based on the corrected classified sentiments. Figure 2 shows the results accuracy of classification testing for existing and proposed method. The plot shows a graph with the results from estimating every classifier's accuracy with the various numbers of features. The accuracy is calculated based on how many numbers of corrected classified sentiments are there out of number of classified sentence correspond to the user's emotion about online product. One of the general purpose machine learning techniques is called maximum entropy classifier which removes the possible least unfair estimation in the given information. It mainly focuses on the elimination of unnecessary information that lies in the reviews or feedbacks[11].

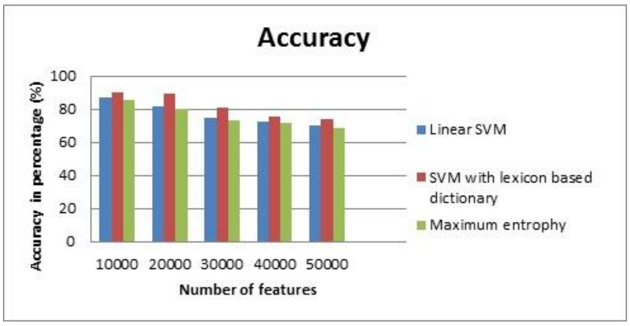

Figure 2 Accuracy of proposed and existing methods

From the above graph, the proposed method has taken high accuracy compared to existing methods Linear SVM and Maximum entropy. From the comparison accuracy results, it is proved that the proposed system can give the high values of corrected classified sentiments. Therefore, the proposed system has the best performance on various numbers of features. The accuracy of each algorithm is the lowest while the total number of features has been utilized.

Table 2 Precision, Recall and F1-Measure of existing and proposed method in the feature size 20000

Metrics	Proposed method SVM with lexicon dictionary approach (%)	Linear SVM (%)	Maximum Entropy (%)
Precision	92.7	88.3	85.4
Recall	90.4	85.9	82.6
F1-Measure	91.53	87.083	83.97

Table 2 shows the precision, recall and F1-measure of existing and proposed methods in the feature size of. In table 2, the proposed take 92.7% of precision, 90.4% of recall and 91.53% of F1-Measure in the feature size 20000. The existing methods Linear SVM has taken 88.3% of precision, 85.9% of recall and 87.083% of F1-Measure, and Maximum entropy has taken 85.4% of precision, 82.6% of recall and 83.97% of F1-measure. Therefore, from the above table, the proposed system has high precision, recall and F1-Measure than the existing methods linear SVM and maximum entropy.

Runtime of Classification

The runtime is the time taken to classify sentiments using the existing and proposed methods has shown in figure 3. From the graph, these classifiers are tested on different size of features 10000, 20000, 30000, 40000 and 50000. The proposed method SVM with lexicon based dictionary approach has taken the shortest time to classify the total online product review dataset, needs 1500ms, 1700ms, 2998ms, 3000 ms in the feature size of 10000, 20000, 30000 and 40000 to preprocess, extract features and classify the sentiments as positive, negative or neutral sentiments on online product review dataset.

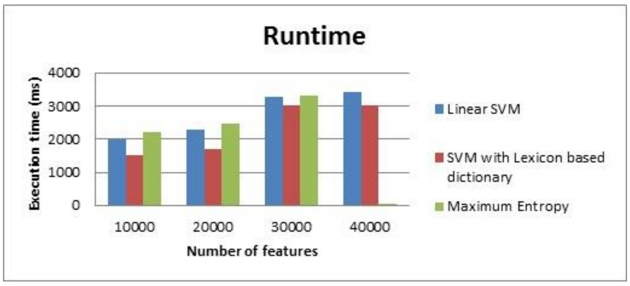

Figure 3 Classification runtime of existing and proposed methods

DOI : 10.18231/2454-9150.2018.1413

Existing methods linear SVM takes time 2000ms, 2300, 3255ms, 3400ms and Maximum entropy method has taken 2200, 2450, 3300, 3350 ms to classify the sentiments in the feature size of 10000, 20000, 30000 and 40000. Hence, from the results proposed system has taken less runtime to classify the sentiments compared to existing methods linear SVM and Maximum entropy.

V. CONCLUSION

In this paper, a new approach of SVM is proposed with lexicon based dictionary for sentiment classification on online product review to classify the user's emotion. The preprocessing method has been executed to filter the unwanted noise and unwanted text from the sample data set. PCA is used to extract the feature efficiently and also it can reduce the multidimensionality complexity. SVM classifier has been employed to classify the sentiments as positive, negative or neutral; here SVM took support of lexicon based dictionary for word matching to classify the sentence. Finally, the dataset of online product review was classified to positive, negative, or neutral effectively. The proposed system is evaluated with respect to the parameters precision, recall and f-measure. It has the result of 92.7, 90.4 and 91.53 respectively which is more accurate when compared to the existing techniques. The runtime results of proposed system have taken lesser time to classify the sentiments compared to existing methods linear SVM and Maximum entropy. By conclusion, the experimental results of proposed system classify the opinions of users accurately with short time.

REFERENCES

[1]. Jinyan Li, Simon Fong, Yan Zhuang and Richard Khoury (2014), "Hierarchical Classification in Text Mining for Sentiment Analysis", International Conference on Soft computing & Machine Intelligence, pp.46-51.

[2]. Kishorjit Nongmeikapam, Dilipkumar Khangembam, Wangkheimayum, Hemkumar, Shinghajit Khuraijam and Sivaji Bandyopadhyay (2014), "Verb Based Manipuri Sentiment Analysis", International Journal on Natural Language Computing (IJNLC) Vol. 3, No.3, pp.113-119.

[3]. Xing Mou and Yajun Du (2016), "Sentiment Classification of Chinese Movie Reviews in Micro-Blog Based on Context", IEEE International Conference on Cloud Computing and Big Data Analysis, pp.313-318.

[4]. Zhang H, Yu Z, XU M and Shi Y (2011), "Feature-level Sentiment Analysis for Chinese Product Reviews", IEEE International Conference on Computer Research and Development, Vol.2, pp. 135 – 140.

[5]. Dan Li, Jiang Qian (2016), "Text Sentiment Analysis Based on Long Short-Term Memory", First IEEE International Conference on Computer Communication and the Internet, pp. 471 – 475.

[6]. Tejasvini Patil and Sachin Patil (2013), "Automatic generation of emotions for social networking websites using text mining", Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Vol.4, pp.1-6.

[7]. Paramesha, K and Ravishankar, K C (2015), "Exploiting dependency relations for sentence level sentiment classification using SVM", IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), pp.1-4.

[8]. Agarwal, A, Xie, B, Vovsha, I, Rambow, O and R. Passonneau (2011), "Sentiment analysis of twitter data," in Proceedings of the Workshop on Languages in Social Media. Association for Computational Linguistics, pp. 30-38.

[9] Zhang, Z, Ye, Q, Zhang, Z and Li, Y (2011), "Sentiment classification of Internet restaurant reviews written in Cantonese," Expert Systems with Applications, vol. 38, no. 6, pp. 7674–7682,

[10]. Chen, J, Huang, H, Tian, S and Qu, Y (2009), "Feature selection for text classification with Naïve Bayes," Expert Systems with Applications, vol. 36, no. 3, pp. 5432–5435.

[11]. Manimekalai S, Santhi A (2016), "Analysis of Big Data Trends and Future Directions: A Review", International Journal of Emerging Technology in Computer Science & Electronics, ISSN: 0976-1353 vol.20, no. 3, pp. 5-8.

 DOI : 10.18231/2454-9150.2018.1413