I J C T A, 9(41) 2016, pp. 677-685 ISSN: 0974-5572

A QUANTITATIVE ASSESSMENT OF MACHINE LEARNING APPROACHES TOWARDS SENTIMENTAL PREDICTION

Sudharshan Duth P and Narendra Kamath G

Abstract: Ample amount of opinion based information is generated on the Internet each and every minute in the form of tweets, reviews or feedback. These texts when taken in small number is possible for a human being to identify the sentiment of the text being positive, negative or objective, but when taken in ample amount, the classification becomes a complicated process. The possible methodology which can be used as a solution for this problem is Sentimental Analysis using machine learning techniques. This paper discusses about the use of various machine learning techniques viz. attribute selection, effective preprocessing, context switchers handling, negation handling and with the quantitative assessment of Naive Bayes, Multinomial Naive Bayes and Sequential Minimal Optimization (SMO) algorithm for the sentimental prediction purpose.

Key Words: *Sentiment Analysis, Naive Bayes, Multinomial Naive Bayes, Sequential Minimal Optimization, Attribute Selection.*

I. INTRODUCTION

The amount of information generated on the Internet is increasing day by day, but there arises a question whether this information is taken into consideration of knowledge extraction or not. Here comes Data mining to the rescue. It focuses on the process of the extraction of valuable data from the information available called as knowledge extraction. The knowledge which is taken into consideration in this paper is the opinion. The opinions are important factors to judge whether there is positive, negative or objective sentiments towards a particular subject taken into consideration. The sentimental analysis is sometimes referred to as opinion mining. Opinion mining is a trending research topic which focuses on the opinion-based textual classification into positive, negative or objective sentiment. The sentimental analysis can be effectively used as an analysis tool for various applications like movie reviews classification, prediction of the winning candidate in the elections based upon the open access information or micro-blogging contents like tweets [19], customer reviews [16] [24] on various products available on online shopping websites [3], student feedbacks which will be taken in the educational sectors [11] for the evaluation of a teacher and much more. The movie reviews from the Internet Movie Database (IMDB) [1] [6] [22] which is widely used as a dataset for the sentimental analysis research work is taken into consideration in this paper. The detailed discussion on the dataset taken into consideration is explicated in further sections.

The opinion based sentimental prediction can be an easy task for a human being to classify whether a given text is positive, negative or objective. Suppose if the individual is given with huge amount of data to classify or if the individual does not know any particular word or its meaning, then it will be a challenging task even for a human being to classify the opinion of the text. In order to train a computer

Deptt. of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Amrita University, Mysuru Campus, Karnataka, India
Emails- sudhisp@gmail.com, narenk53@gmail.com

program to understand the human opinions based on the text is another challenging task which can be addressed by sentimental analysis as discussed earlier. This paper discusses on the solution towards effective sentimental analysis and prediction by using several machine learning techniques viz. attribute selection, effective preprocessing, context switchers handling, negation handling along with the inclusion of the quantitative assessment of the classification algorithms which are widely used for sentimental analysis viz. Naive Bayes, Multinomial Naive Bayes and Sequential Minimal Optimization (SMO). There is an observation made that with the implementation of the machine learning techniques as discussed earlier, there is a significant improvement in the accuracy of the sentimental classification process. In the further sections there is a detailed discussion of the machine learning techniques implemented in this paper.

II. SENTIMENT CLASSIFICATION AND ITS TYPES

Sentiment classification is an emerging field of study under the main domain of Data Sciences. It deals with the processing of the input data being in the form of image, text or sound and with the extraction of the opinion being positive, negative or objective, based on the provided input [2] [20]. The output of the sentimental classification can be used for the analysis purpose which can be called as sentimental analysis. There is a lot of data available in the form of text on the Internet, which is expressed by human beings in order to exchange the opinions in the form of reviews, feedbacks or micro-contents, which makes it suitable for the implementation of sentimental classification and analysis on the textual content. Sentiment classification can be broadly categorized into three types, they are as follows,

2.1 Sentence Level Sentiment Classification–

In this category of sentiment classification, the text will be taken as the input in the form of a single sentence. This particular sentence will be subject to the pre-classification, classification and the post-classification steps in order to fetch the opinion out of it. It is one of the simplest, yet very important categories to be taken into consideration. Some works have done the simplest of having to match the positive and negative words and then to decide onto the opinion being expressed in the sentence. But this is not a proper way to compute the opinion, since there requires the contextual polarities to be computed in order to fetch the opinion.

2.2 Document Level Sentiment Classification–

There might be an assumption made that the document level of sentiment classification is just the extension of sentence level of sentiment classification. However, it is just a misconception made on the matter. The document level of sentiment classification when considered to be an aggregate of sentences will take high amount of resources for the computation of the sentiment. When the number of documents considered is less, it will be feasible to implement the extension of the sentence level sentiment classification/analysis. But, when the number of documents is more, then the deployment of the suitable classification algorithms for the document level sentiment classification will be beneficial to serve the purpose.

2.3 Feature Level Sentiment Classification–

As the number of attributes increases for the sentimental computation, the complexity of the sentiment classification increases. As the name implies, the sentiment of a text is determined by the feature which is taken into consideration. This type of sentiment classification is beneficial for the contextual based sentimental analysis. The feature set which is considered, is usually ranked [25] and restricted, because the less and the most appropriate features when taken into consideration, will be beneficial towards the betterment of the sentimental classification process.

The accuracy for the above mentioned categories of sentiment classification can be improvised with the addition of machine learning techniques viz. feature selection/ attribute selection, negation handling, effective pre-processing and any suitable classification algorithm based upon the type of the sentiment classification taken into consideration.

III. RELATED WORK

There is a lot of research being done on the topic of opinion mining and sentimental analysis. Some of the papers concentrate on the discussion or the survey being done on the topic of sentimental analysis [20]. Whereas some papers focuses on the provision of a guide towards the sentimental classification of the text. But many of which, focus on the improvement to be made on the sentimental classification part, with the inclusion of the methodologies which supplement towards the improvement of sentimental classification accuracy. Bo Pang et al. [1] have used the movie reviews as the dataset and has also examined about applying machine learning methods for the sentiment classification problem and there is a mention of SVM outperforming Naive Bayes for the sentiment classification problem. Bo Pang et al. [2] have covered the techniques and the approaches towards the provision of opinion-oriented information retrieval systems. The work [2] also focuses on the material on briefing of the evaluative text and on the broader issues with respect to the privacy, manipulation and economic effect on the development of opinion based information seeking systems. Rohini K et al. [3] have proposed the opinion miner system, which is supposed to be designed to mine the customer opinions [16] based on the product reviews. The paper [3] uses the dataset of Amazon and other publicly available dataset for the purpose of experimentation. Stefano Baccianella et al. [4] have explained about Sentiwordnet 3.0, which is a lexical resource that is publicly available for research purposes. The Sentiwordnet 3.0 is used in this very paper for the purpose of determining polarity of specific words in the initial stages of experimentation. Trivedi Khushboo et al. [5] have proposed the methodology to classify the text at the sentence level with the use of the Naive Bayes algorithm for classification. Vivek Narayanan et al. [6] have used the Naive Bayes algorithm for the sentimental classification of IMDB movie reviews with fast and accurate Naive Bayes with the negation handling and feature selection. Bo Pang et al. [7] have proposed a novel methodology of machine learning that applies text categorization techniques for the subjective portions of the document. In this paper [7], there is a mention of the efficient techniques for finding minimum cuts in graphs that can be applied; which greatly facilitates incorporation of cross-sentence contextual constraints. Liu et al. [8] have defined the opinion mining problem. The paper [8] describes various key mining tasks that have been studied in their research literature and their representative techniques. There is a discussion about the issue of detecting opinion spam or fake reviews. Finally, there is a mention of the research topic of assessing the utility or quality of online reviews. Abhishek Tiwari et al. [9] have proposed the TF-IDF (Term Frequency-Inverse Document Frequency) methodology for opinion mining process, where the same was tested on the real time data from various websites and blogs and 81% accuracy was achieved on the same. Hoeber et al. [10] paper focuses on the individual tweet sentimental classification which is based on the visual analytics technique called as Vista.

Dhanalakshmi V et al. [11] paper focuses on the opinion mining process on the dataset being considered as student feedbacks. The paper [11] focuses on the various classification algorithms like SVM, Naive Bayes, K-Nearest neighbor and neural network classifier being deployed for opinion mining process with the comparison of the same algorithms taken into consideration. Pablo Gamallo et al. [12] have described a strategy based on a Naive-Bayes classifier for detecting the polarity of English tweets. In the paper [12], there is a mention of the classifier being deployed only for two sharp polarities being positive and negative, where in which, is stated to be factors towards improvement in performance. There are also some methodologies which are implemented using the subjectivity analysis or the domain-adaptation analysis as discussed by Gezici, Gizem et al. [13] which has also proved to be a feature which can be implemented for sentiment analysis, which was implemented using Weka tool. The semantic and sentimental similarities between the words can also be captured using unsupervised and extension towards the supervised models for the purpose of prediction of the sentimental observations in the context where the words actually appear by learning the word vectors, as discussed by Andrew L Maas et al. [14]. Arti Buche et al. [15] have surveyed and analyzed the various techniques that have been developed for the primary tasks of opinion mining. There is a mention of an overall picture of what is involved in expanding a software system for opinion mining on the basis of their survey and analysis. Dr. Ritu Sindhu et al. [16] have tried to demonstrate the clustering and classifying opinion mining experimentation on analysis of web blog posts on recent product policy and services reviews. The paper [16] focuses on a novel approach for analyzing the customer opinions.

IV. METHODOLOGY

The sentimental classification in this paper is implemented using various machine learning techniques. The input dataset which is taken into consideration in this paper consists of 1000 positive reviews and 1000 negative reviews of the Internet Movie Database (IMDB), which was introduced in Pang/Lee ACL 2004, released June 2004. The IMDB movie reviews dataset which is widely used for the purpose of sentiment analysis [1] [6] [20] [22], gave a motivational factor to use the same for this work. The methodologies implemented in this very paper are as discussed below.

4.1 Effective Pre-processing–

Pre-processing is considered as an important step to be carried out in a data mining process. In this paper there is an implementation of the effective pre-processing technique. This process includes the following steps:

- The first step is to remove the extra white spaces and extra tab spaces.
- In the next step the URLs will be removed.
- In consideration with the current texts available, even the hash tags will be removed. But when taken tweets [19] into consideration this will not be an optimal method of pre-processing, since hash tags also include some subjective/sentimental information.
- The next step is to remove the non-alphabetic characters except the single quotation mark. This is because the negation words like can't, don't and other similar words will be having the single quotation mark which will be useful for the ease in identification of negation words in the dataset.
- The stop words are removed in the next step. This is done by not considering a standard list of stop words but with the consideration of the stop words list custom built with no subjective words included. For this process the SentiWordNet [4] was referred in order to get the subjective nature of a word.
- The last step is to transform the text to a common lower case.

4.2 Negation Handling–

The pre-processed text is then given as input to the next module of negation handling, where if a negation word like not, n't words like doesn't, can't, won't etc. occurs, then from the next word, all the words will be prefixed with NOT_ [6] such that the words will be negated with their polarities. This process is continued till a context switching word like but, however, although etc. occurs. The process of negation handling is done in order to match to the real-time scenario texts [21] where there will be completely a positive sentence but the negation word will change the complete opinion itself. E.g. the sentence 'He is not a good boy' will be negated as 'He is not NOT_a NOT_good NOT_Boy' when taken as in the context of unigram.

4.3 Context Switchers–

This step is carried out in parallel to the previous step i.e. negation handling, where in which there arises a situation where the half of the sentence will give an opinion but the other half will give the contrasting opinion. So, in that scenario there is a requirement to shift the context of the sentence. The context here refers to the opinion which is fetched. E.g. the sentence 'He is not a good boy, but he is an excellent boy' will be transformed to 'He is not NOT_a NOT_good NOT_boy but he is an excellent boy' such that the negation handling is discontinued when context switcher occurs and the context is switched to contrasting opinion of the previously fetched opinion.

4.4 Attribute Selection–

The attribute selection is a process which is done in order to reduce the feature space dimensionality. But before performing the actual attribute selection, the pre-processed documents are merged together and transformed into a single file which will ease the process. The file generated is then transformed to word vectors [14] with minimum term frequency of 3. After the application of the filter, the attribute selection is done with the evaluator as information gain attribute evaluator and with the search model as a ranker with the threshold of zero. In this paper before the attribute selection, there were 2400 attributes which were limited during the vector formation step. After attribute selection the top most ranked 496 attributes were selected and a filtered file was created.

The file size before and after the application of the pre-processing steps and attribute selection is as tabulated in the Table 1.

Table -1Document Size before and after Pre-processing and Attribute selection

Stages	File Size Before	File Size After
Pre-Processing (Inclusive of Negation Handling and Context Switching)	7.54 MB	5.16 MB
Attribute Selection	5.16 MB	821 KB

There is an observation made that after the preprocessing and attribute selection, the file size has reduced in significant amount as tabulated in Table 1. The decrease in the file size is due to the removal of unwanted features/ attributes. This is a key step towards further classification process.

4.5 Sentiment Classification–

The major step in the whole process is the sentimental classification. In this step the classification algorithms Naive Bayes, Multinomial Naive Bayes and Sequential Minimal Optimization are used for the experimentation purpose in this paper. The detailed discussion on the implementation of these algorithms is, as mentioned below.

4.5.1 Naive Bayes– The Naive Bayes classification algorithm is widely used for the sentimental classification purpose. This algorithm is a simple yet powerful probabilistic model with a base of Bayes rule and conditional independence assumption. In this algorithm the text is classified based on the posterior probabilistic value of the class to which it belongs to. The implementation of Naive bayes is using the standard formulation of the same. It is observed that the standard Naive Bayes gave good result of the classification accuracy. The observation is made that the pre-processing, negation handling, context switching and attribute selection methods have supplimented for the improvement in the accuracy of the Naive Bayes classifier. The result achieved using this algorithm is dicussed in the results section of this paper.

4.5.2 Multinomial Naive Bayes– This model is a specialized flavor of Naive Bayes which is said to be a direct generalization of the Binomial Naive Bayes model. This model, unlike simple Naive Bayes is designed especially for the document based text classification. This variation considers the number of occurences of a word in training documents from the respective class, may be positive or negative, with the inclusion of multiple occurences. In this paper the multinomial Naive Bayes was implemented using the standard multinomial Naive Bayes algorithm. The accuracy out of this model was far better than the simple Naive Bayes model considered earlier.

4.5.3 Sequential Minimal Optimization (SMO)– It is an algorithm which is generally used for solving the quadratic programming problem that arises during training of Support Vector Machines (SVMs) [17]

[18]. It is an iterative algorithm in nature which will split the large, complicated problem into a series of smaller problem, which makes it easy for the classification process [17]. It is observed that the implementation of SVM along with the intesifiers and with the modification of the values returned from the lexical datasource Sentiwordnet [23], the accuracy of sentiment classification is improved. But, the major demerit of SVM is that for a large amount of data, the accuracy will be considerably reduced. SMO is a solution for this problem. The implementation of SMO is done using the training dataset being the pre-processed document with the attribute selection too given as input. There was a slight improvement noticed when compared with the multinomial Naive Bayes' accuracy measure which was implemented.

The algorithms which are implemented in this paper were evaluated with a realistic classification evaluation model taken into consideration. For this purpose of evaluation, the 5-fold cross validation was used.

V. RESULTS

The implementation of the classification algorithms for the purpose of sentimental classification yielded good results. The dataset which was used is the 2000 IMDB movie reviews as discussed earlier. The result obtained for each of the algorithms is discussed in the sub-sections.

5.1 Result of simple Naive Bayes model–

The Naive Bayes model when implemented on the data yielded 76.6% accuracy. The result obtained from the implementation of the simple Naive Bayes model is as tabulated in Table 2.

Table -2 Result of simple Naive Bayes model

Sl No.	Parameter Name	Result
1.	Correctly Classified Instances	1532
2.	Incorrectly Classified Instances	468
3.	Kappa statistic	0.532
4.	Mean absolute error	0.2342
5.	Root mean squared error	0.4789
6.	Relative absolute error	46.8416 %
7.	Root relative squared error	95.7732 %
8.	Total Number of Instances	2000

In Table 2 the observation can be made that the simple Naive Bayes algorithm yielded 76.6% accuracy with f-measure score of 0.7494.

5.2 Result of Multinomial Naive Bayes model–

The implementation of the Multinomial Naive Bayes model yielded 84.35% accuracy. The result obtained from the Multinomial Naive Bayes model implementation is as tabulated in Table 3.

Table -3Result of Multinomial Naive Bayes model

Sl No.	Parameter Name	Result
1.	Correctly Classified Instances	1687
2.	Incorrectly Classified Instances	313
3.	Kappa statistic	0.687
4.	Mean absolute error	0.1589
5.	Root mean squared error	0.3635

6.	Relative absolute error	31.7789 %
7.	Root relative squared error	72.7095 %
8.	Total Number of Instances	2000

In Table 3 the observation can be made that the multinomial Naive Bayes model yielded 84.35% accuracy with f-measure score of 0.8437.

5.3 Result of Sequential Minimal Optimization model–

The implementation of the Sequential Minimal Optimization (SMO) model yielded 86.4% accuracy. The result obtained from the SMO model implementation is as tabulated in Table 4.

Table -4Result of Sequential Minimal Optimization (SMO) model

Sl No.	Parameter Name	Result
1.	Correctly Classified Instances	1728
2.	Incorrectly Classified Instances	272
3.	Kappa statistic	0.728
4.	Mean absolute error	0.136
5.	Root mean squared error	0.3688
6.	Relative absolute error	27.2 %
7.	Root relative squared error	73.7564 %
8.	Total Number of Instances	2000

In Table 4 the observation can be made that the SMO model yielded 86.4% accuracy which is a slight improvement seen when compared with the multinomial Naive Bayes model implemented earlier.

These classification algorithms were implemented on the processed information as stated in earlier sections. The 5-fold cross validation model was included for the evaluation purpose. The metrics considered for the evaluation of these models are as tabulated in Table 5.

Table -5Metric Evaluation of Classification Algorithms Implemented

Metrics	Classification Algorithms		
	Simple Naive Bayes	***Multinomial Naive Bayes***	***SMO***
Accuracy	76.6%	84.35%	86.4%
Precision	0.8064	0.8424	0.8691
Recall	0.7	0.845	0.857
F-Measure	0.7494	0.8437	0.8630

For the ease of the visualization purpose the graphical representation of the accuracies of the implemented classification algorithms is as depicted in Figure 1.

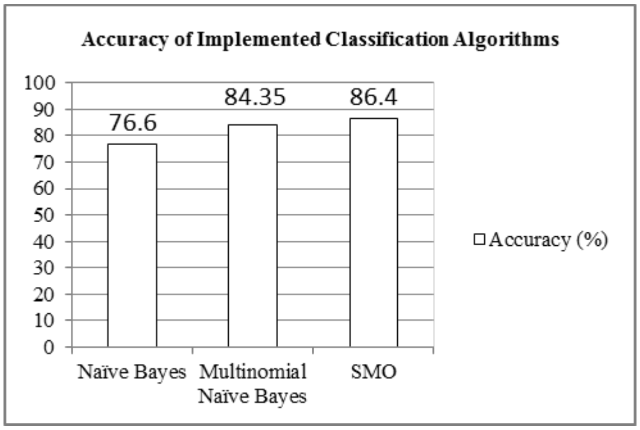

Figure 1. Comparison of the accuracy of various classification algorithms implemented

The results obtained after the application of the sentimental classification algorithms show that the initial process of pre-processing and the attribute selection have constituted towards the improvement of the sentimental classification process and it is observed that SMO model has outperformed the other two models with 86.4% accuracy.

VI. CONCLUSION AND FUTURE WORK

Extraction of the sentiment or the opinion from the textual content is an important aspect in data analytics and advanced text processing, where in which the classification algorithms should work well to serve the purpose of sentimental classification. The opinion mining task is implemented in this paper with the consideration of the three classification algorithms, viz. the simple Naive Bayes model, the Multinomial Naive Bayes model and the Sequential Minimal Optimization model. It is observed that with the inclusion of machine learning techniques like effective preprocessing, negation handling, context switchers handling and attribute selection, the accuracy of the classification algorithms have improved significantly. The observation is made that the SMO model outperformed the other two models with a higher accuracy level of 86.4%, with the restricted number of feature set taken into consideration. It is noticed that the Multinomial Naive Bayes model's accuracy was close to the accuracy level of SMO model, which shows that the Multinomial Naive Bayes model can be improvised in order to achieve significantly higher accuracy levels. It is also noticed that the sentimental classification methodology when applied to an individual domain taken into consideration, the accuracy level will be more. Mapping to the real world scenario, the sarcastic text handling can be beneficial towards the overall accuracy improvement of the sentimental analysis in the near future. The ideas discussed in this paper can be used as optimal solution towards the development of sentimental analysis applications.

REFERENCES

[1] Bo Pang and Lillian Lee ShivakumarVaithyanathan, "Thumbs up?Sentiment Classification using Machine Learning Techniques," Conference on Empirical Methods in Natural Language Processing (EMNLP),Volume 10, pp. 79-86, 2002.

[2] Bo Pang, "Opinion Mining and Sentiment Analysis," 2008.

[3] Jin, Wei, Hung Hay Ho and Rohini K. Srihari, "OpinionMiner: a novel machine learning system for web opinion mining and extraction." Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM,pp. 1195-1204, 2009.

[4] Stefano Baccianella, Andrea Esuli and FabrizioSebastiani, "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," In Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10), Vol. 10, pp. 2200-2204, 2010.

[5] Khushboo T N, Vekariya S K & Mishra S, "Mining of Sentence Level Opinion Using Supervised Term Weighted Approach of Naïve Bayesian Algorithm," International Journal of Computer Technology and Applications, 3(3), 2012.

[6] Narayanan V, Arora I & Bhatia A, "Fast and accurate sentiment classification using an enhanced Naive Bayes model," In International Conference on Intelligent Data Engineering and Automated Learning, Springer Berlin Heidelberg, pp. 194-201, 2013.

[7] Bo Pang and L Lee , "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," In Procedings of the ACL, pp. 271, 2004.

[8] Liu B & Zhang L, "A survey of opinion mining and sentiment analysis," In Mining text data, Springer US, pp. 415-463, 2012.

[9] AbhishekTiwari, KshitijPathak, Upasanatiwari and Rupam das, "Opinion Polarity Detection in Blog Comments from Blog Rss Feed by Modified TF-IDF Algorithm," International Journal of Engineering Research and Applications (IJERA), Vol. 2, pp. 412-416, 2012.

[10] O Hoeber, L Hoeber, L Wood, R Snelgrove, I Hugel, and D Wagner, "Visual Twitter analytics: Exploring fan and organizer sentiment during Le Tour de France," In Proceedings of the VIS 2013 Workshop on Sports Data Visualization, pp. 1-7, 2013.

[11] Dhanalakshmi V, Bino D and Saravanan A M, "Opinion mining from student feedback data using supervised learning algorithms," In2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), pp. 1-5, 2016.

[12] Pablo Gamallo and Marcos Garcia, "A Naive-Bayes Strategy for Sentiment Analysis on English Tweets," Proceedings of the 8th International Workshop on Semantic Evaluation, pp. 171–175, 2014.

[13] Gezici, Gizem, Berrin Yanikoglu, Dilek Tapucu, and Yücel Saygın, "Sentiment Analysis Using Domain-Adaptation and Sentence-Based Analysis," In Advances in Social Media Analysis, pp. 45-64. Springer International Publishing, 2015.

[14] Andrew L Maas, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng and Christopher Potts, "Learning Word Vectors for Sentiment Analysis," The 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011),Volume 1, pp. 142-150, 2011.

[15] Arti Buche, Dr. M B Chandak and Akshay Zadgaonkar, "Opinion Mining and Analysis: A Survey," International Journal on Natural Language Computing (IJNLC), Vol. 2, 2013.

[16] Dr. Ritu Sindhu, Ravendra Ratan Singh Jandail and Rakesh Ranjan Kumar, "A Novel Approach for Sentiment Analysis and Opinion Mining," International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 4, 2014.

[17] Keerthi, S Sathiya, Shirish Krishnaj Shevade, Chiranjib Bhattacharyya and Karuturi Radha Krishna Murthy, "Improvements to Platt's SMO algorithm for SVM classifier design," Neural Computation, 13(3), pp.637-649, 2001.

[18] Thelwall, Mike, Kevan Buckley, Georgios Paltoglou, Di Cai and Arvid Kappas, "Sentiment strength detection in short informal text," Journal of the American Society for Information Science and Technology, 61(12), pp.2544-2558, 2010.

[19] Pak, Alexander, and Patrick Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," In LREc, vol. 10, pp. 1320-1326. 2010.

[20] Medhat, Walaa, Ahmed Hassan and Hoda Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, 5(4), pp.1093-1113, 2014.

[21] Dadvar, Maral, Claudia Hauff and F M G De Jong, "Scope of negation detection in sentiment analysis," 2011.

[22] Sharma, Anuj and Shubhamoy Dey, "Performance investigation of feature selection methods and sentiment lexicons for sentiment analysis," IJCA Special Issue on Advanced Computing and Communication Technologies for HPC Applications, 3, pp.15-20, 2012.

[23] Bhaskar, Jasmine, K Sruthi and Prema Nedungadi, "Enhanced sentiment analysis of informal textual communication in social media by considering objective words and intensifiers," In Recent Advances and Innovations in Engineering (ICRAIE), 2014, pp. 1-6. IEEE, 2014.

[24] Gangothri V, S Saranya and D Venkataraman, "Engender Product Ranking and Recommendation Using Customer Feedback," In Proceedings of the International Conference on Soft Computing Systems, pp. 851-859. Springer India, 2016.

[25] Rajeev, P Venkata and V Smrithi Rekha, "Opinion Mining of User Reviews Using Machine Learning Techniques and Ranking of Products Based on Features," In Proceedings of the International Conference on Soft Computing Systems, pp. 627-637. Springer India, 2016.