

# APSO BASED FEATURE EXTRACTION APPROACH AND SVM BASED CLASSIFICATION FOR SENTIMENTAL ANALYSIS

KUMAR ATTANGUDI PERICHIAPPAN PERICHAPPAN<sup>\*</sup>,  
SREENIVAS SASUBILLI<sup>\*\*</sup>, ABHISHEK KUMAR<sup>\*\*\*</sup>

## ABSTRACT

The focal point of Sentimental Analysis is intolerance of estimation or else some sentiment so as to survive in the assessment. Based on the three categories, the execution of Sentiment Analysis has been done such as several supervised, un-supervised and machine learning algorithm. Nevertheless, each and every model has its own advantages as well as disadvantages. An effort has been given by our above mentioned proposed methodology to establish partial swarm intelligence based sentimental supervised methodology. Since huge quantity of data samples in sort to attain an appropriate characteristic data set, the above mentioned model has been used to achieve the highest finest characteristic set. Minimum Redundancy as well as Maximum Relevancy is used to calculate the evaluation of finest attribute set is achieved. By utilizing with the taxonomy of Support Vector Machine, the cataloging of the excerpt feature set has been proficient. The calculation of recall, precision, accuracy and f-measure are taken here for analysis.

## INTRODUCTION

Sentimental Analysis is referred as to recognize the sentiment of consumers i.e., to know the individual person what he/she is saw, to understand the meaning which is said by the people and finally have to know what they says and it is also called as opinion mining [17]. In recent times or else the recent workplace, they have to understand the above mentioned emotions in software side as well as it is essential to identify with the producers and the director of business. By using natural language

processing (NLP), we can easily realize the social media communications. Here, it will be categorized into three polarities such as positive, negative and neutral. If the communication is based on positive, then it is called as positive polarity, if it is based on negative, then we can say it as negative polarity and the combination of both positive and negative content, then it is said to be neutral [28]. And the major vital role of sentimental analysis is accuracy.

---

\* Associate Director KPMG, Regional Development Center Roseland, NJ, USA.

\*\* Solution Consultant, KPMG, Regional Development Center Roseland, NJ, USA.

\*\*\* Assistant Professor, Computer Science Engineering Department, ACERC, Ajmer.

**Correspondence E-mail Id:** editor@eurekajournals.com

The applications of sentimental analysis are followed by, support in decision making, business application and predictions and trend analysis [1]. Furthermore, it will be categorized into two divisions such as (i) Lexicon analysis which is used to estimate the divergence in the communication in social media commencing with the keywords of semantic direction as well

as the perspective will not be considered for applications which is based on lexicon analysis and (ii) Machine learning (ML) is used to identify the orientation of the keyword from the labelled training dataset which consists of building models [15]. For further references [2], [6], [7], [8] and [14]

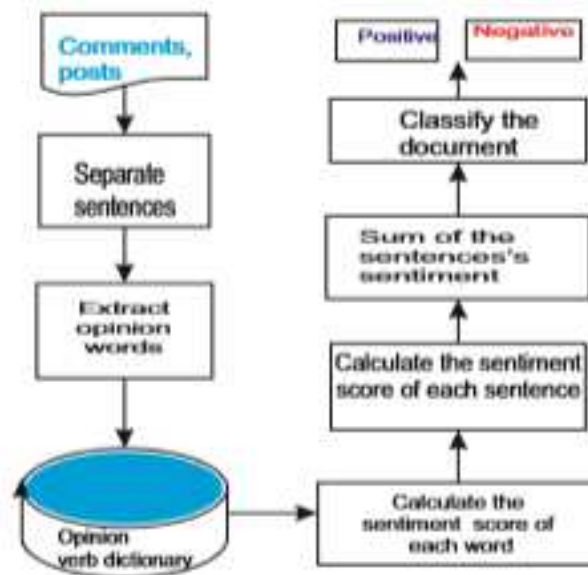


Figure 1. Architecture of Sentiment analysis [16]

## MACHINE LEARNING IN SENTIMENTAL ANALYSIS

An anthology of dataset which consists of labelled tweets is referred as machine learning for sentimental analysis. Due to the noisy information, we can easily pre-process the data by utilizing several Natural Language processing (NLP) methods. Moreover, the related characteristics for sentiment analysis are must to be extorted. Finally, it will be skilled as well as experienced on concealed information [3]. Furthermore, the following algorithms are stronger at the time of come together along with auto-tagging algorithms namely, LDA, Auto-Tag URL and Named Entity Recognition. Based on sentiment classification, the following algorithms are used for machine learning algorithm such as, Naive Bayes classifier, Support vector machines and decision trees.

**NAIVE BAYES CLASSIFIER**-Most of the researchers used this type of algorithm and it will be effortless method. To envisage the characteristic possibilities, we used to refer this classifier.

**SUPPORT VECTOR MACHINES**-In explore space, it is used to detect linear separators which is used to divide the various modules. It is aptly for text categorization.

**DECISION TREES**-The skilled information space will be symbolized in a hierarchical framework because to detachment the data, featured charge has been used.

## CHALLENGES OF SENTIMENT ANALYSIS

There are various levels are produced in someone verdict which is left over to the above mentioned problem. To understand the above mentioned dilemma, there is merely one

method through framework which has to be acquainted with an initial stage and it could be powerfully brunt the emotion of presently interior verdict.

At last, to skill the model which one has to be choose is the additional challenge in sentiment analysis.

**PRE-PROCESSING**

By using the framework of have another look at which the informations are offered through the users as it consists of heavy noise has to be eliminated preceding towards it to be classified. Enhance the exactness of the classification system by using assortment of appropriate characteristics. For skilled characteristics, unigrams and bigrams have been chosen which is used to enhance the accuracy [16].

**FEATURE EXTRACTION**

It is under from feature based sentiment analysis. It is used to identify the characteristic of products which has been mentioned by the consumers. Feature extraction refers that it took text as input as well as produces the extorted characteristics in one among the following forms such as Lexico-Syntactic or Stylistic, Syntactic and Discourse based. Furthermore, it will be classified into various types and some of them are as follows: (i) Morphological types, (ii) frequent features and (iii) implicit features [5].

**CLASSIFICATION**

The block diagram for the classification of sentimental analysis is given below as,

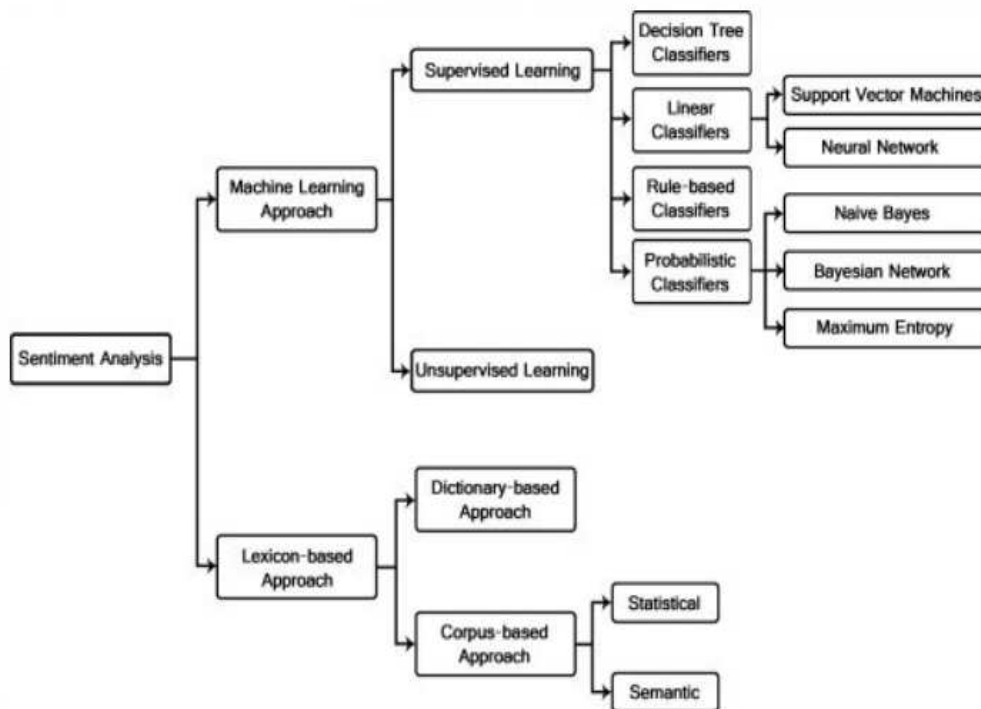


Figure 2. Classification of sentiment analysis [29]

**LITERATURE SURVEY**

Harshali P. Patil and Dr. Mohammad Atique have taken a survey about Sentiment Analysis for Social media: a survey [17]. Nowadays, social media such as facebook, Twitter, etc, has

become the resolution creating system for any type of social activities and people used to share their sentiments and thoughts through the social media. Sentiment analysis for social media has been analyzed for such activities which are above mentioned. The only target is

to recognize the data in the Web as well as it must be categorized based on the polarities of sentiment analysis-positive, negative and neutral. Finally, based on various methods, the survey has to be taken to recognize the type of work.

Lukasz Culer and Olgierd Unold developed a technique based on Sentiment Analysis of social networks statements for the Polish language [18]. Here, the authors have the intention of the above mentioned research is used to acclimatize a sentiment analysis system and the declaration of social network is used to attain the sentiment range in the Polish Language. By considering the neighbourhood of sentiment, we can easily identify the bigram sentiment through the initialization of PMI-IR. The bigram is used to imply the performance is high. Finally it will be compared with the previous generated algorithms, which is concluded in less effective solutions.

Samhaa R. El-Beltagy and Ahmed Ali proposed a system as Open Issues in the Sentiment Analysis of Arabic Social media: a case study [19]. Based on the various micro-blogging intermediates will become insist for Arabic sentiment analysis equipment due to the enhancement in the quantity of Arabic supporting situations. By comparing the Arabic sentiment analysis along with English as well as other languages, there will be a low systematic for Arabic analysis. Here, the authors presented a major problems and issues in the above mentioned study as well as to examine the probability of identifying the semantic course of Arabic resource has been established through a case study.

Giulio Angiani, Laura Ferrari, Tomaso Fontanini, Paolo Fornacciari, Eleonora Iotti, Federico Magliani and Stefano Manicardi introduced a model based on A comparison between Pre-processing Techniques for Sentiment Analysis in Twitter [4]. Pre-processing is the initial step in

sentiment analysis but it will not be fully focussed on literature. Here, the authors focussed on pre-processing system and it is used to enhance the accuracy of the method. The following parts are discussed and presented in the proposed system such as dissimilar methods, significance and accuracy. The major intention of this paper is to estimate the models which will be more effectual. Furthermore, the proposed method will be compared to the previous existing models.

Emma Haddi, Xiaohui Liu and Yong Shi developed a system based on The Role of Text Pre-processing in Sentiment Analysis [20]. Nowadays, it's become more confront to recognize the new tendency of social media as well as the size will be high and the essential of authentic period estimation removal along with mining. Most of researches will not concentrate on pre-processing methods. So, the authors investigated about pre-processing method in sentiment analysis and presented a description based on feature selection by utilizing support vector machines (SVM). Through the SVM, we can easily increase the accuracy of the system. The performance of the proposed system is much better when compared to others.

Savitha Mathapati, S H Manjula and Venugopal K R introduced a model based on Sentiment Analysis and Opinion Mining from Social media: a review [16]. As we discussed in preceding papers, most of the researches will not concentrated on pre-processing method of sentiment analysis. Here, the authors introduced a technique based on 3 sections in sentiment analysis such as pre-processing, feature extraction and classifications. Pre-processing technique is used to increase the accuracy of the system. The important factor in sentiment analysis is pre-processing method which is used to get the results exactly. In this pre-processing method, unigrams and bigrams are used to skill the characteristics of the system. Finally, the performance of the system

is much better when compared to the previous existing systems.

M. Trupthi, Suresh Pabboju and G. Narasimha proposed a technique based on Improved Feature Extraction and Classification-Sentiment Analysis [12]. The authors discussed the proposed method by comparing with the preceding existing system. Here, to overcome all the problems in the existing system the authors introduced a machine learning classifications to achieve the feature extraction model will be more effective and efficient one. Since, the following categories are based on feature based sentiment analysis such as, feature extraction, sentiment classification and sentiment evaluation. The sentiment analysis model has been achieved for movie reconsider datasets. The performance of the proposed method is compared with the existing system and the result will be much better.

Farzana Kabir Ahmad introduced a model based on Comparative Analysis of Feature Extraction Techniques for Event Detection from News Channels' Facebook Page [22]. Here, the authors has been taken comparison the proposed method to the existing system and several feature extraction models has been proposed. In this proposed method, three various classifiers are discussed such as Support vector machine (SVM), Naive Bayes (NB) and K-Nearest Neighbor (kNN). The target of the research is to introduce the feature extraction model and it is used to identify the event which provides finest accuracy end result. The proposed model will be checked based on the following channels such as BBC, Aljazeera and Al-Arabiya. The performance of the proposed method is much better when compared to others.

Rajalaxmi Hegde and Dr. Seema S developed a model based on Aspect Based Feature Extraction and Sentiment Classification of Review Data sets using Incremental Machine

learning Algorithm [23]. To detect the sentiments as well as phase is the mainly vital dispute in sentiment analysis. The process will perform based on data classifications and so called aspect based feature mining. Using machine learning, the summarization has to be done based on the following factors such as Support vector machine (SVM), Naive Bayes (NB), maximum entropy and finally Random forest technique. Finally, the performance will be compared with the existing system and all the problems have been overcome in the proposed system as well as the performance is much better.

Kirti Huda, Md Tabrez Nafis and Neshat Karim Shaukat suggested a system based on Classification Technique for Sentiment Analysis of Twitter Data [24]. The behavior of the user can be considered as sentiment analysis as well as the twitter data will be considered as data. In sentiment analysis, four steps have to be taken. The first step is pre-processing which is used to improve the accuracy of the system. The second step is feature extraction which is used to identify the characteristics from the customers. The final step is classification of sentiment analysis. Here, the second step is taken as input. Moreover, the prototype model has been pertained to the feature extraction in which the prototype models are produced since existing systems are used to enhance the accurateness of the classification of data. By using nltk tool box the proposed system has been executed in python. Finally, the implementation time is decreased as well as accuracy has been enhanced.

Ajay Deshwal and Sudhir Kumar Sharma proposed a model depends on Twitter Sentiment Analysis using Various Classification Algorithms [25]. Wherever in the world, we can easily detect the happened activities by using Twitter via micro blogging feature. It will be small as well as create regularly by the people. Based on the polarity of sentiment analysis, the

messages have been classified such as positive, negative and neutral. Here, the authors proposed a model which is used to merge the feature extraction as well as classification of sentiment analysis which is used for more accuracy in the system. The performance is good when compared with the existing system of the model. Furthermore, the comparison will be done based on the six organized classification algorithms.

Ms. K. Mouthami, Ms. K. Nirmala Devi and Dr. V. Murali Bhaskaran introduced a technique based on Sentiment Analysis and Classification Based on Textual Reviews [13]. To use social media, the features of frame of mind or emotion of customers has been analyzed routinely and it is used to establish the polarities of consumer taking place in the web. Document-level sentiment classification is used for articulate a positive or negative sentiment. Here, the accuracy of classification is less and we can consider the sentiment analysis by using multi-theme document is little complex. By using Bag of words in SVM, we can able to order the document level. To enhance the classification accurateness, a trendy algorithm has been introduced along with parts of speech label as Sentiment Fuzzy Classification. Finally, the performance is good.

## **PROPOSED WORK**

### **PROPOSED PARTIAL SWARM INTELLIGENCE BASED FEATURE SELECTION AND SUPPORT VECTOR MACHINE BASED CLASSIFICATION APPROACH**

An Intelligent Supervised Semantic Analysis Approach is established by way of utilizing the optimization algorithm designed for capable characteristic excerption since huge data illustrations has been discussed in the below part. As we have notified in the literature survey, here survive frequent methods intended for taxonomy as well as collecting the

semantic information commencing the huge information illustrations. Furthermore, to perform an accurateness as well as efficient model, the proposed system is fully focused on the excerpt of similar stipulations or else words since the semantic data have been discussed in this part. The figure 2 represents the architecture diagram of the proposed model. Moreover, for efficient feature excerpt, a developmental technique has been utilized so called Particle Swarm Optimization. Based on three sections, the proposed model has been classified such as:

- Pre-processing Module
- Feature Extraction Module
- Classification Module

#### **PRE-PROCESSING MODULE**

The pre-processing module is used to achieve or else collect the data to transmit the collected informations into a suitable form of natural language processing to accomplish a qualitative sentiment analysis as well as the steps are to be followed as:

- Process English language: The performance of testing as well as calculation only on English language has been proposed in our model and it will eliminate the data which consists of other languages. For semantic analysis WordNet6 dictionary is used which is having only in English language.
- Stemming and lemmatization: The process of minimizing the words to its root is called stemming and the process of spelling modification to the stemming term is called lemmatization that will avoid the unpredictability of words at the time of vocabulary accomplishment.
- Spelling correction: To achieve the performance will be much better, there is a necessity of spelling correction and it will be attained through JSpell.

- Removal of stop words, query words, URLs, special characters: For sentiment organization, the factors will not be supplied routinely. For furthermore process, there will be a necessity of stop words as well as query words.
- Enlarge abbreviations and Substitute jargons: There is a necessity to so to facilitate whichever irregularity is circumvented at the time of vocabulary proficient.
- Part of speech tagging: At the time of sentiment detection is used, there will be a necessity of part of speech knowledge. Under observation, the part of speech study will be more important which uses Stanford POS tagger to distinguish the information from part of speech.

## FEATURE SELECTION MODULE

### STEPS FOR FEATURE SELECTION USING PSO:

- i. Initialization of population: The candidate solutions or a particle is defined as by choosing the pre-processed informations from the size of the individuals or swarm has been initialized which is used for categorization beside with the target labels which has been allocated to every swarm.
- ii. Initializing the term location and velocity: By using a vector, each and every factor has a position in the investigation area which is given as vector  $x_i = (x_{i1}, x_{i2}, \dots \dots x_{iD})$   
Where,  
  
D-Magnitude of exploration domain.  
  
To accomplish the outcome here we use personality meander in the investigation area. So, each and every personality has a speed, which is indicated as  $v_i = (v_{i1}, v_{i2}, \dots \dots v_{iD})$ .
- iii. Fitness Evaluation: To sort out the similar factors since the pre-processed information sample, Minimum Redundancy Maximum

Relevancy (MRMR) has been used to be calculated. In this section, the particles are taken as robustness of factors in PSO algorithm. With the absence of association among the characteristics, we can use the peak ranking characteristics which are based on communal data as MRMR.

## CLASSIFICATION MODULE

For additional categories of the information samples, from the feature selection module, the characteristic has been attained. The following polarities are classified from the class labels or else objective class of the information sample such as completely negative polarity, negative polarity, neutral, positive polarity and completely positive polarity. Based on two sets, the entire data has been separated such as training set as well as testing set in which the characteristics of testing samples utilize the Particle Swarm Optimization Algorithm commencing the feature selection module. To create a categorization module, it will utilize the training data set along with class labels as well as to organize the manuscript; it will utilize the testing data samples which are connected to the characteristic factors.

For the classification of information sample, we can utilize the SVM. The whole numeric vectors in domain have been plotted by SVM as well as denote that result restrictions via hyper planes. The vector has been separated in two categories by hyper plane as well as hyper plane is maximum for each and every category of distance.

## EXPERIMENTAL RESULTS

By employing two irrelevant information sets, the performance of the proposed system has been displayed. Through the use of TripAdvisor, we can calculate the proposed system along with tiny or no noise in textual form for hotel evaluation. By employing the following

categories, the model concludes as well as it will be matched with the predictable classification of sentimental using SVM as Sentimental Classification along with the pre-processed module and SVM model. To calculate the efficiency of proposed method, the four dissimilar measures are given below as:

Precision: it is also known as positive predicted value. The suitable portions of recovered illustration have been measured by precision which is denoted as:

$$precision = \frac{TP}{TP+FP}$$

Recall: The section of suitable illustrations is recovered from recall as well as it is also known as sensitivity and it is denoted as:

$$recall = \frac{TP}{TP+FN}$$

Accuracy: It is defined as the common proportion of accurate illustrations which is irrespective of its kind. By using the accuracy, we can calculate the efficiency of sentiment analysis which is denoted as:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

Where

TP-true positive

TN-true negative

FP-false positive and

FN-false negative

F-Measure: It will permit factoring accuracy repeatedly as well as it will have the capability to recollect into an exceptional range by indicating its harmonic mean. And it is denoted in the given formula:

$$F - measure = \frac{(1+\beta^2) \times recall \times precision}{(\beta^2 \times recall) + precision}$$

Where

$\beta$ -used to identify higher weight in the score or to recall.

$\beta = 1$  is used to represent the balancing load to the two metric.

## RESULT ANALYSIS

Step 1-**importdata**-loading the dataset into the program

Step 2-**Preprocessing**-we need to load the TA entire data into training and testing database which are done based on the preprocessing that is stemming process.

Step 3-**Feature extraction**-> here we have to apply partial swarm optimization algorithm

In this, we are finding the BestSol.Cost, now we have to **multiply this optimal cost with the training and testing data.**

This plot represents the cost is gradually reducing while increasing the iteration, after sometimes which attains the constant or convergence stage.

Step 4-Support Vector machine classification-final step is the classification, here we have to train and test the database which is getting from feature extraction using Partial Swarm optimization.

Step 5-performance evaluation-now we need to evaluate performance based on following parameters are Accuracy, precision, Recall and F-measure. These parameters are evaluated between Output\_Labels and test\_labels (it is actual labels)

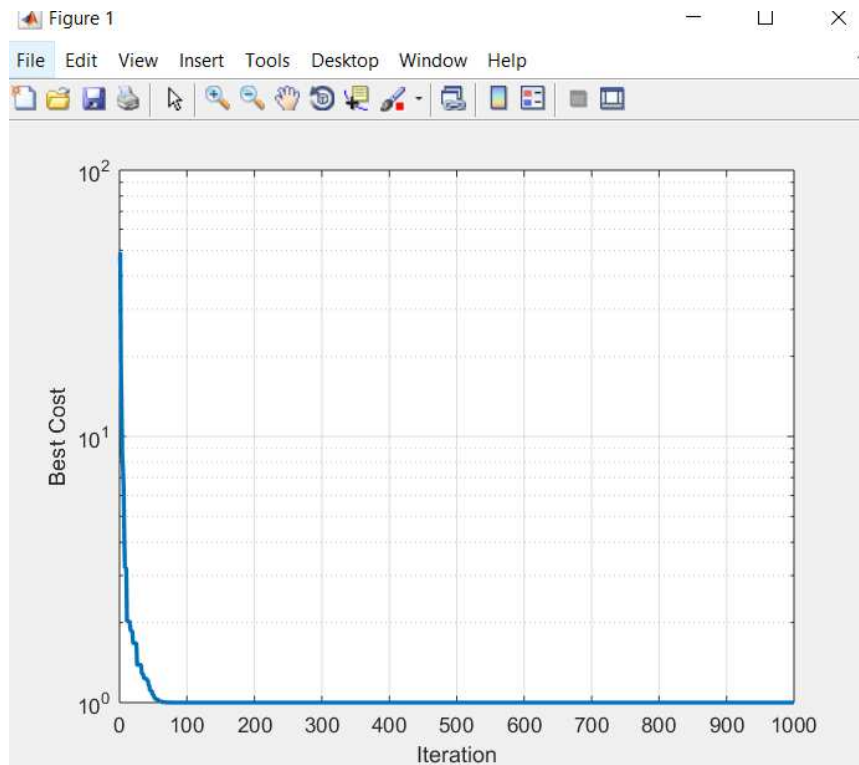
Accuracy-86.2489

Precision-72.4345

Recall-100

F\_measure-84.0139





## CONCLUSION

Supervised sentiment analysis conclusion investigation showings fair execution levels anyway it needs ordered corpora which are not easily reachable for every locale. Here, to achieve the accuracy, a swarm intelligence based sentimental classification algorithm is used and it is compared with the existing sentimental classification algorithms. Since the huge data samples, the primary concentration has been given for the suitable chosen of characteristics to estimate the factors in the informations in terms of calculation of fitness. For the estimation of minimum redundancy as well as maximum relevancy, the above mentioned fitness calculation has been used in the particle swarm optimization. At first, the pre-processing of the manuscript is proficient pursued pertinent element choice and grouping utilizing support vector machine. By using precision, recall, f-measure and accuracy, we can concede the conclusion of the proposed model. The performance of the proposed model is much better and it will be

demonstrated when compared along with the further two preceding models.

## REFERENCES

- [1]. Harpeet Kaur, Veenu Mangat and Nidhi, "A Survey of Sentiment Analysis techniques", published in International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 978-1-5090-3243-3/17/\$31.00 ©2017 IEEE, (I-SMAC 2017).
- [2]. Oscar Romero Llombart, "Using Machine Learning Techniques for Sentiment Analysis", published in June of 2017, School of Engineering (UAB).
- [3]. Anuja P Jain and Asst. Prof Padma Dandannavar, "Application of Machine Learning Techniques to Sentiment Analysis", published in 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 978-1-5090-2399-8/16/\$31.00 c 2016 IEEE.
- [4]. Giulio Angiani, Laura Ferrari, Tomaso Fontanini, Paolo Fornacciari, Eleonora

- lotti, Federico Magliani, and Stefano Manicardi, "A Comparison between Pre-processing Techniques for Sentiment Analysis in Twitter".
- [5]. Muhammad Zubair Asghar, Aurangzeb Khan, Shakeel Ahmad and Fazal Masud Kundi, "A Review of Feature Extraction in Sentiment Analysis", published in Journal of Basic and Applied Scientific Research, 4(3)181-186, 2014 © 2014, TextRoad Publication.
- [6]. Pietro Ducange and Michela Fazzolari, "Social Sensing and Sentiment Analysis: Using Social Media as Useful Information Source", 978-1-5386-2101-1/17/\$31.00 ©2017 IEEE.
- [7]. Sola Hong, Jaedong Lee and Jee-Hyong Lee, "Competitive Self-Training Technique for Sentiment Analysis in Mass Social Media", published in SCIS&ISIS 2014, Kitakyushu, Japan, December 3-6, 2014, 978-1-4799-5955-6/14/\$31.00 ©2014 IEEE.
- [8]. Anurag D. Kshirsagar and Dr. Kalpana S. Thakre, "A Survey on Analyzing Social Media Data for Understanding Students' Learning Experiences", published in International Journal of Computer Science and Mobile Computing, Vol.5 Issue.6, June- 2016, pg. 76-82.
- [9]. Zhao Jianqiang and Gui Xiaolin, "Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis", DOI 10.1109/ACCESS.2017.2672677, IEEE Access.
- [10]. Akrivi Krouska, Christos Troussas and Maria Virvou, "The effect of pre-processing techniques on Twitter Sentiment Analysis", 2010.
- [11]. Zhao jianqiang, "Pre-processing Boosting Twitter Sentiment Analysis?", published in 2015 IEEE International Conference on SmartCity/SocialCom/SustainComtogether with DataCom 2015 and SC2 2015.
- [12]. M. Trupthi, Suresh Pabboju and G.Narasimha, "Improved Feature Extraction and Classification-Sentiment Analysis", published in International Conference on Advances in Human Machine Interaction (HMI-2016), March 03-05, 2016, R. L. Jalappa Institute of Technology, Doddaballapur, Bangalore, India.
- [13]. Ms.K.Mouthami, Ms.K.Nirmala Devi and Dr.V.Murali Bhaskaran, "Sentiment Analysis and Classification Based On Textual Reviews", 2013.
- [14]. Oskar Ahlgren, "Research on Sentiment Analysis: The First Decade", 2014.
- [15]. Imane El Alaoui<sup>1</sup>, Youssef Gahi, Rochdi Messoussi, Youness Chaabi, Alexis Todokoff and Abdessamad Kobi, "A novel adaptable approach for sentiment analysis on big social data", published in J Big Data (2018) 5:12 <https://doi.org/10.1186/s40537-018-0120-0>.
- [16]. Savitha Mathapati, S H Manjula & Venugopal K R, "Sentiment Analysis and Opinion Mining from Social Media: A Review", published in Global Journal of Computer Science and Technology: C Software & Data Engineering, Volume 16, Issue 5, Version 1.0 Year 2016.
- [17]. Harshali P. Patil and Dr. Mohammad Atique, "Sentiment Analysis for Social Media: A Survey", published in 978-1-4673-8611-1/15/\$31.00 ©2015 IEEE.
- [18]. Lukasz Culer and Olgierd Unold, "Sentiment analysis of social networks statements for the Polish language", published in 2016 Third European Network Intelligence Conference, 978-1-5090-3455-0/16 \$31.00 © 2016 IEEE DOI 10.1109/ENIC.2016.26.
- [19]. Samhaa R. El-Beltagy and Ahmed Ali, "Open Issues in the Sentiment Analysis of Arabic Social Media: A Case Study", 2015.
- [20]. Emma Haddi, Xiaohui Liu and Yong Shi,

- "The Role of Text Pre-processing in Sentiment Analysis", published in 2013 International Conference on Information Technology and Quantitative Management doi: 10.1016/j.procs.2013.05.005.
- [21]. R. A. S. C. Jayasanka, M. D. T. Madhushani, E. R. Marcus, I. A. A. U. Aberathne and S. C. Premaratne, "Sentiment Analysis for Social Media", published in Conference Paper November 2013, online available in <https://www.researchgate.net/publication/268817500>.
- [22]. Farzana Kabir Ahmad, "Comparative Analysis of Feature Extraction Techniques for Event Detection from News Channels' Facebook Page", published in Journal of Telecommunication, Electronic and Computer Engineering, e-ISSN: 2289-8131 Vol. 9 No. 1-2.
- [23]. Rajalaxmi Hegde and Dr. Seema. S, "Aspect Based Feature Extraction and Sentiment Classification of Review Data sets using Incremental Machine learning Algorithm", published in 3rd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB17).
- [24]. Kirti Huda, Md Tabrez Nafis and Neshat Karim Shaukat, "Classification Technique for Sentiment Analysis of Twitter Data", published in International Journal of Advanced Research in Computer Science, Volume 8, No. 5, May-June 2017, Available Online at [www.ijarcs.info](http://www.ijarcs.info).
- [25]. Ajay Deshwal and Sudhir Kumar Sharma, "Twitter Sentiment Analysis using Various Classification Algorithms", published in 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Sep. 7-9, 2016, AIIT, Amity University Uttar Pradesh, Noida, India.
- [26]. Walaa Medhat, Ahmed Hassan and Hoda Korashy, "Sentiment Analysis algorithms and applications: A survey", published in Ain Shams Engineering Journal (2014) 5, 1093-1113.
- [27]. Youngsub Han and Kwangmi Ko Kim, "Sentiment Analysis on Social Media Using Morphological Sentence Pattern Model", published in 978-1-5090-5756-6/17/\$31.00 ©2017 IEEE SERA 2017, June 7-9, 2017, London, UK.
- [28]. Kumar, A., & SAIRAM, T. (2018). Machine Learning Approach for User Accounts Identification with Unwanted Information and data. International Journal of Machine Learning and Networked Collaborative Engineering, 2(03), 119-127.
- [29]. Rawat K., Kumar A., Gautam A.K. (2014) Lower Bound on Naïve Bayes Classifier Accuracy in Case of Noisy Data. In: Babu B. et al. (eds) Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012. Advances in Intelligent Systems and Computing, vol 236. Springer, New Delhi DOI: [https://10.1007/978-81-322-1602-5\\_68](https://10.1007/978-81-322-1602-5_68).