SENTIMENT ANALYSIS USING MACHINE LEARNING

Krishma
Dept. Computer Engg. &Technology (M. Tech)
Guru Nanak Dev University (Amritsar)

Amit Chhabra
Dept. Computer Engg. & Technology
Guru Nanak Dev University (Amritsar)

***Abstract*- In this paper, we discussed Sentiment analysis. In past years, the web has huge amount of data on social media sites. Social media such as Twitter, Facebook etc. where billions of peoples expressed their own emotions i.e. sad and happy emotions. In other words, we can say Internet is growing day by day. Facts and opinions are two main parts of Textual information. In facts information, only subjective information are available and in opinion information represents the sentiments of their presenters. In this paper we discuss introduction of Sentiment analysis, different methods of Sentiment analysis, Sentiment classification techniques and Machine learning algorithms for Sentiment classification.**

Keywords- Sentiment Analyzer, Twitter, Machine Learning Techniques

I. INTRODUCTION

On the Internet textual information is growing every day. Search and analysis the textual information on Internet are very difficult to day by day. Now a days, web search engines retrieves the different types of information or data from the web-based providers i.e. Social sites, searching sites etc. for example, if we try the word "nature" on searching sites like Google, yahoo or other searching sites then it searches different meaning of "nature". In this paper we discussed Twitter social media site [5]. Twitter has only 280 characters limit which are approximately a newspaper headlines or subheading of something that have restricted length of characters.

[13] There are different approaches for sentiment analysis: Machine learning based approach (ML) uses many machine learning algorithms (supervised or unsupervised algorithms) to classify the data. Lexicon based approach uses positive and negative words that are determined the polarity of words. And last one is Hybrid approach, in which combination of both lexicon-based approach and Machine learning approach.

Sentiment analysis the process of computationally identifying and categorizing opinion expressed in form of text, especially a topic, product etc. that are positive, negative or neural. Sentimental analysis are process of automatically identifying whether a user-generated text expresses positive(+ve), negative(-ve) or neutral opinion about an entity (i.e. products, peoples, topics, event etc). Sentiment classification can be done at Document level, Sentence level, aspect or Feature level. In Document level the whole documents are used as a basic information unit to classify it either into positive or negative classes.[8] Document level determine the overall opinion of product this document level very useful if he documents are related only a one entity e.g. person, product, place etc. Sentence level sentiment classification classifies each sentence first as subjective or objective and then classifies into positive, negative or neural class.[8] Sentence level consists each sentence as a separate unit. Aspect or Feature level or entity level sentiment classification deals with identifying and extracting product features from the source data [10]. Aspect level mixture of emotion from the review sentence. The main concept or goal of aspect is to find the sentiment on entities and their aspects [8]. In bags of word, relationship between the individual words are not considered. In Sentiment score, functions of scoring apply on text or reviews of the tweets in Twitter.

Sentiment analysis study is to study the opinions of people and their emotional views that are stated in the form of text. Organizations use sentiment analysis to monitor their business to find out customer's mindset and also to find out what customer's think about their services. This could be a major uplift for their business enhancement. Fig.1 represents the procedure involved in sentimental analysis. Sentiment analysis or opinion mining can be categorized into three points: document level, sentence level, entity and aspect level. The steps explained of a flowchart, as below: -

- Firstly, in the flowchart, we set the goal and scope of text content.
- Secondly, cleaning data in which unusable data are removal. When we taking data from web then more different type of data are available like structured, unstructured or semi-structured. So, clean it must.
- Then, comes bag of the words in which reliable and useful data are stored.
- In last step score function is performed. Scoring is a process in which data or content are analyzed.

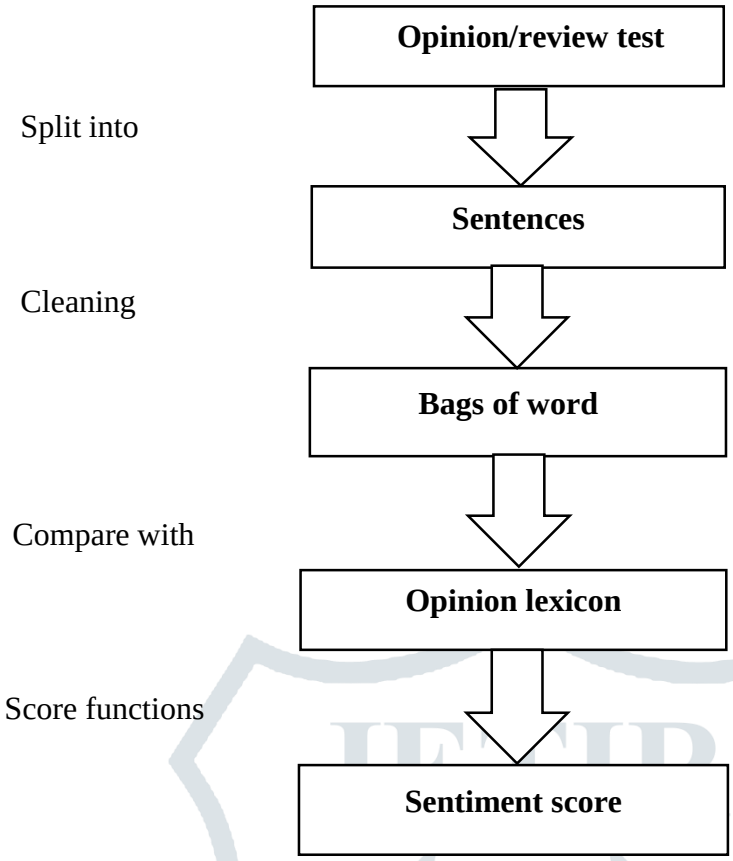


Fig 1. Sentiment analysis process

Machine learning is most effective method used in the field of data analysis in order to predict something by devising some models and algorithms. These analytics models allow researchers, engineers, data scientists and analysts to produce reliable and valid results and decisions. Feature selection is most important task of machine learning. Machine learning used in those problems where theoretical knowledge is insufficient and using this, we have numbers of observation and final outputs or we can say final results.

II. LITERATURE REVIEW

Penubaka Balaji et al. in 2018[1] An overview on opinion mining techniques, in this studied several opinion mining techniques and sentiment analysis techniques and their accuracy in the opinion or sentiment categories. In this paper, sentiment analysis process, opinion/sentiment components and types.

Flora Poecze et al. in 2018[2] Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts, the Support vector learning method "k nearest neighbor" was chosen for the Sentiment classification of the retrieved commentaries. In this paper analysis the Facebook metrices. The ANOVA of post subcategories and their Facebook metric means demonstrated that photos depicting the family, friends, pets compared to subcategories with the lowest means in the sample.

Ali Hasan et al. in 2018[3] Machine learning based Sentiment Analysis for Twitter Accounts, in this paper author's include the adoption approach of Hybrid that involves a sentiment analyzer that including machine learning. In this paper framework of tweet classification and sentiment analysis are also discussed.

M.Trupthi et al. in 2017[4] Sentiment Analysis on Twitter using streaming API, In this paper provides an interactive automatic system, posted tweets by people in social media using Hadoop in which huge amount of data. This paper main objective to perform sentiment analysis on the tweets that are extracted from the twitter.

Sheena Angra et al. in 2017[5] Machine learning & its application, described the machine learning's application. machine learning techniques includes k-nearest neighbor, multilayer perceptron, self-organization maps (SOM), and statistical techniques and the results of these techniques. in this paper applications of machine learning are discussed with their relevant examples.

S. Rajalakshmi, S. Asha et al. in Survey on sentiment analysis. This paper keywords used like opinion mining, sentiment analysis and social media. In this paper process of sentiment discussed. In this paper applications of sentiment analysis also used.

Harpreet Kaur, Veenu Mangat in 2017[7] A survey of Sentiment analysis techniques, describe in this paper applications of sentiment analysis, sentiment analysis process and classification techniques.

Shivaprasad TK, Jyothi Shetty in 2017[8] Sentiment analysis of product reviews: A Review, in this paper discussed levels-based sentiment analysis in which document level, sentence level and aspect level. In this paper sentiment analysis process discuss and classification techniques used for accuracy.

Neha Raghuvanshi et al. in 2016[9] A brief review of Sentiment analysis, presents new avenues in Opinion mining and Sentiment Analysis. The findings in this paper are as follows. Opinion mining will be specific to each user's or group of user's preferences and needs.

Neha Raghuvanshi, Prof. J.M patil 2016[10] A brief review on Sentiment analysis discussed in this paper naïve bayes classifier used. In this paper different methods of used to mine the data on the web. On the web day to day data increased. We have proposed sentiment analysis based on Naïve Bayes.

Fatehjeet Kaur Chopra et al. in 2016[11], in this paper different levels of classification in sentiment analysis are discussed and different levels of classification are: - document level, sentence level and aspect level. All references to the given subject are detected by SA, and sentiment in each reference using natural language processing (NLP) techniques is also determined by SA.

Melody Moh et al. in 2015[12] On Multi-Tier Sentiment analysis using supervised machine learning. it includes a multi-tier prediction model, which applies various supervised machine learning methods to predict sentiment levels. Four classifiers, Naïve Bayes, supervised vector machine, Random forest and Stochastic gradient descent have been applied to this model.

Jonnathan Carvalho et al. in 2014[14] A Statistical & Evolutionary Approach to sentiment Analysis, in this paper classification of tweets in which extracting tweets feature, computing and polarity of each feature and computing the polarity final.in this paper researchers describe the genetic algorithm for word selection problem. Genetic analysis used initial population, reproduction and new generations and stop.

Anuja P Jain et al. in 2016[15], in this paper researchers says that twitter is a type of social media site, that have huge amount of data and fast emergent micro-blogging social networking platform for expressed the user expressions, emotions etc. In this paper application of Machine Learning are discussed.

Aliza Sarlan et al.in 2014[16] Twitter sentiment analysis is developed to analyze customers perspectives towards the critical to success in the marketplace. The program is using a machine-based learning approach which is more accurate for analyzing a sentiment; together with natural language processing techniques will be used.

Neethu M S et al. in 2013[17] Sentiment Analysis in twitter using Machine Learning, there are different symbolic and machine learning techniques to identify sentiments from text. Classification precision of the feature vector is verified using different classifiers like Nave Bayes, Maximum Entropy, SVM.

Afroze Ibrahim Baqapuri in 2012[19] Twitter Sentiment Analysis, discussed in this paper, classification, feature extraction, TweetMod Web applications, In Tweet Mood web application, Tweet score, TweetCompare and TweetStats are used. Sentiment analysis especially in the domain of micro-blogging.

Arundhati Navada, Aamir Nizam Ansari et al. in 2011[20] Overview of use of Decision tree algorithms in Machine Learning, in this paper machine learning, decision tree, ID3, IDA, C4.5 are used. In this paper comparison between IDA and ID3 also discussed.

Following terms are used in this paper also: -

A. Opinion Mining

Opinion mining refers the large area of language processing, text mining, sentiments, opinions and emotions expressed in text. Although, view or attitude based on emotion instead of reason is often referred to as a sentiment. Opinion mining is a platform that is used for collect the emotions and know about what people think and feel about their product and services [18].

B. Twitter

Twitter is a popular social media site that permit users to share short information or knowledge that called tweets which are limited to 280 characters. Users write tweets to direct their view about several topics relating to their daily survives. Twitter is an ideal platform for the extraction of general public opinion on specific issues. A group of tweets is used as the main body for sentiment analysis, which refers to the use of opinion mining or natural language processing.

C. Twitter Sentiment Analysis

The sentiment can be found in the comments or tweet to provide useful indicators for many different purposes. Also, and stated that a sentiment can be categorized into two groups, which is negative and positive words. Sentiment analysis is a natural language processing technique to quantify an expressed opinion or sentiment within a selection of tweets [14].

III. DIFFERENT METHODS IN SENTIMENT ANALYSIS

Sentiment analysis mainly three techniques Supervised learning, Unsupervised learning and Hybrid. These are following discussed: -

A. Supervised Learning Algorithms

Supervised Learning techniques in Sentiment Analysis have been SVM, Nave Bayesian Classifiers and other Decision Trees. [7] SVM is a sophisticated approach in which only labeled training documents. Maximum Entropy is conditional exponential classifier. ME maps feature of each pair set and its label to a vector. It is loglinear classifier because they work by extracting some set of features from the input [4]. In this learning techniques decision tree, linear, rule based and probabilistic classifier are used.

B. Unsupervised Learning Algorithms

Unsupervised Learning are also called as lexicon-based techniques. This algorithm used to discover the opinion lexicon. Unsupervised learning involves learning patterns in the input when no specific outputs are specified or supplied. In K-means, center of the cluster are finds by calculating the distance from center of cluster, and tries to finds the natural cluster in dataset. In k-means, the positions of cluster are minimal. The centers are randomly assigned in cluster. Unsupervised learning basically two types: Dictionary based approach and Corpus based approach. Dictionary based approach searches the antonym and synonyms words in the dictionary. [7] Other side corpus-based approach starts with opinion words list and find some or more other opinion words.

C. Hybrid Techniques

In Hybrid techniques both combination of machine learning and lexicon base approaches used. Researchers have proved that this combination gives improved performance of classification. The main advantages of their hybrid approach using a lexicon/learning association is to get the best of both worlds-stability as well as readability from a carefully designed lexicon and the high accuracy from a Supervised learning algorithm [4].

In fig-2, different types of Sentiment Analysis. Basically, Sentiment analysis are two types. First one is Machine Learning approach and second one is Lexicon based approach. Further these two types more detailed into more types. Supervised learning and unsupervised learning are types of machine learning and other side dictionary based and corpus based both are types of lexicon based. In Supervised learning decision tree, linear classifier and more classifiers approaches. [18] Decision tree is a hierarchical form in which internal nodes can be taken as tests and leaf nodes can be takes as categories. Decision tree used for various fields. One of the oldest decision tree algorithms is ID3. In lexicon-based approach two types are available. Dictionary based approach searches the antonym and synonyms words in the dictionary and other side corpus-based approach starts with opinion words list and find some or more other opinion words. Classification techniques of Sentiment or Opinion Analysis are in below fig: -

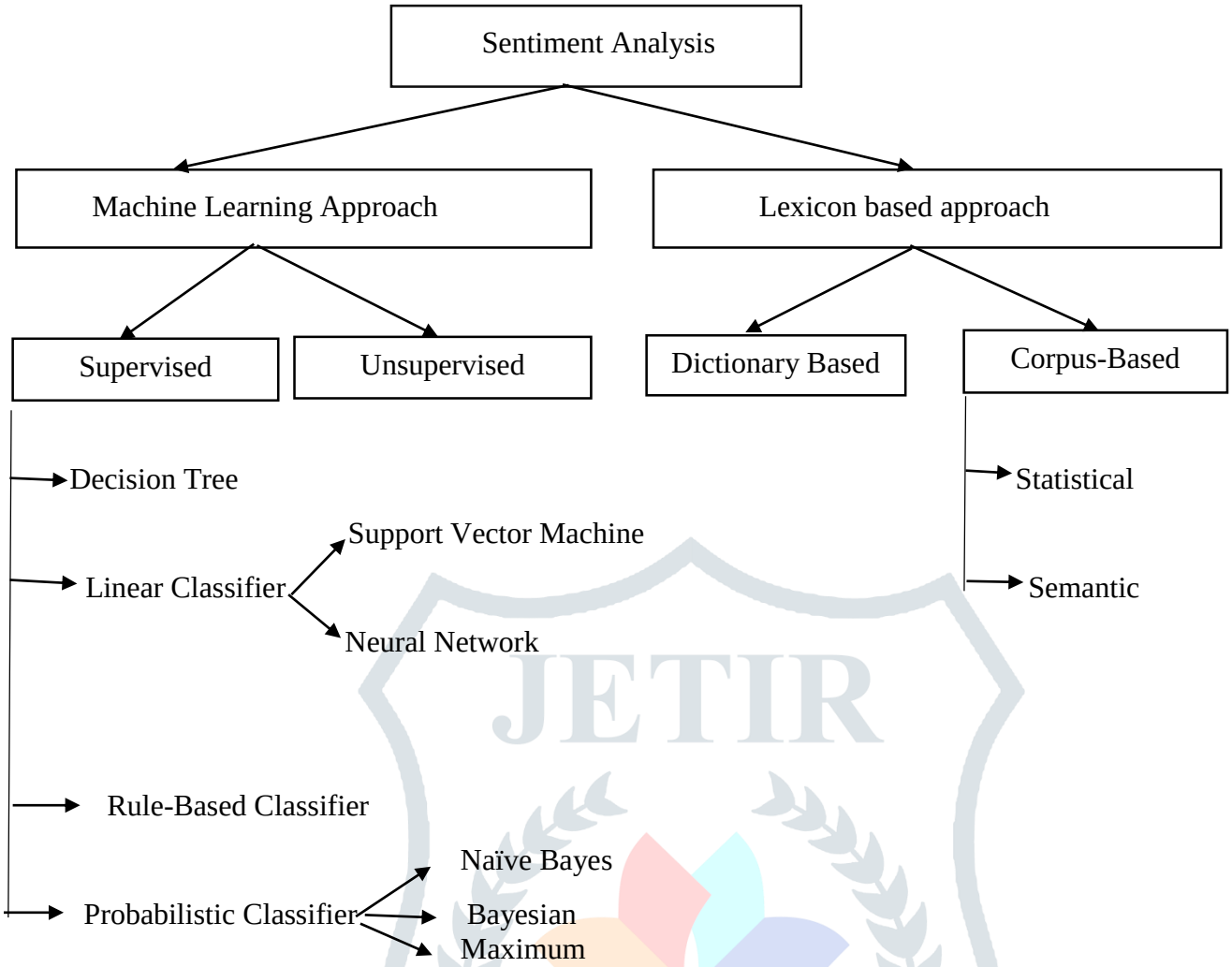


Fig 2: - Sentiment or Opinion Classification Techniques

IV. SENTIMENT CLASSIFICATION TECHNIQUES

A. Lexicon based approach

Lexicon based methods are predefined list of words where each and every word are associated with a specific sentiment or opinion [9]. Finding the opinion lexicon for calculate the sentiment for a given text. It deals with no's of +Ve and –ve words count in the text. By calculating if text consists more no's of +Ve words then it assigned positive score and if the text contains both +Ve and –ve numbers are equally then it assigned neutral score.

B. Machine Learning based approach

[9] Machine learning approaches often rely on supervised classification approaches where sentiment recognition is enclosed as a binary which are positive and negative. Here, two sets of documents are needed: training and a test set. A supervised learning classifier uses the training set to learn and train itself with respect to the differentiating attributes of text dataset. Machine learning have several types of techniques like Support Vector Machine (SVM), Maximum entropy (ME), Naïve Bayes (NB) are used by classify reviews and tweets. Machine learning for sentiment analysis starts with collection of datasets containing labelled tweets.

C. Hybrid approach

To improve Sentiment classification performances some researchers, techniques and algorithms are suggested us using combination of both lexicon based and machine learning techniques. The main advantage of this hybrid approach is that we can attain best of both worlds. The lexicon learning combination has proven to improve accuracy. Lexicon based approach have high precision and low recall. Hence combining it with a machine learning classifier can improve the recall and accuracy of the algorithm [9].

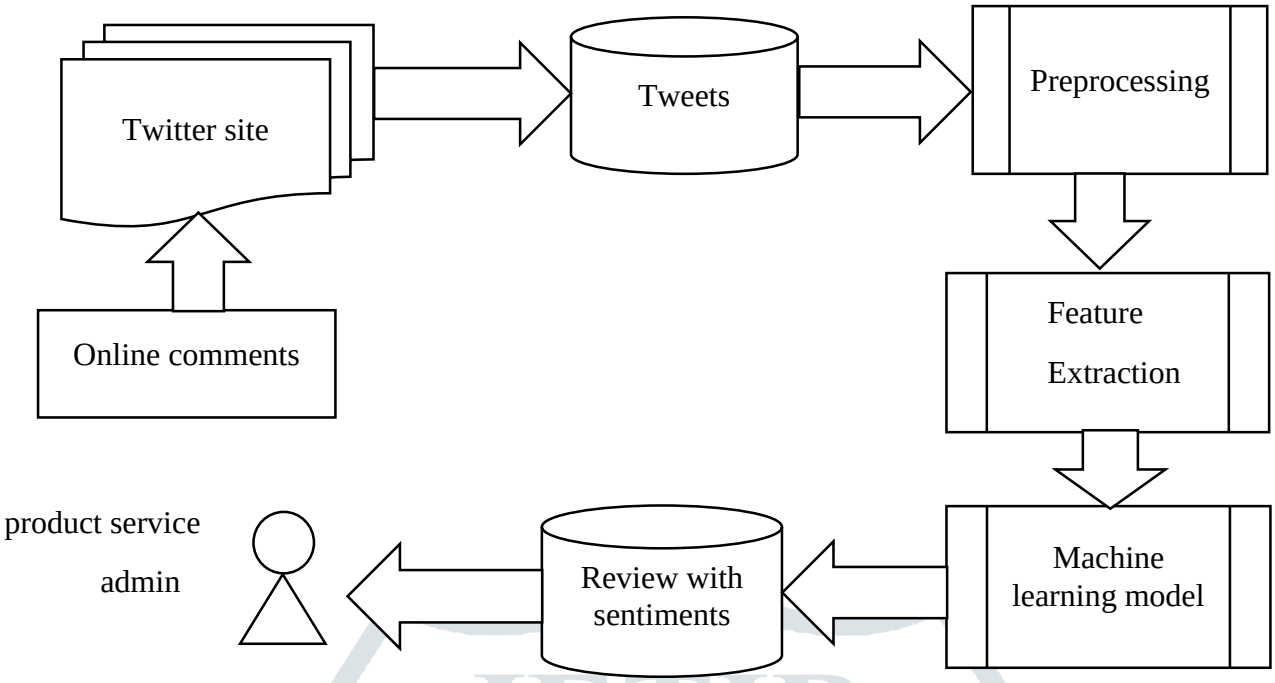


Fig 3: Workflow of twitter sentiment analysis using machine learning algorithms.

In fig-3, workflow of twitter sentiment analysis starts from online comments. These comments are noisy data from different servers or sets. Through twitter site these noisy data in the form of tweets in twitter social site. Now these tweets need to Pre-processing because pre-processing used various natural language processing. Data preprocessing is a process that collects different data from different social sites, online shopping sites, online reservation sites etc. these data may be either structured and unstructured.so like this unwanted information removed at this stage. Next step in workflow is feature extraction. [7] In this phase extract the feature from the text which are N grams, POS tagging, stemming, stop words, conjunction handling etc. then machine learning model have different classification techniques or algorithms that are used for classification.

V. TWITTER SENTIMENT ANALYSIS PROCEDURE

Twitter is a very famous or we can say popular social media for exchange the short messages i.e called Tweets. [2] Twitter is a microblogging site where multiple users exchange messages. Before we perform on twitter the sentiment analysis data the data should be brought into proper form and sentiment relevant features need to be extracted and preprocessing.

The steps followed in twitter sentiment analysis as shown in figure 3 are:

1. ***Data collection****:* Twitter allows Tweets by using a Twitter API. A twitter account has a must to obtain twitter credentials (i.e. API key, API secret, Access token secret) which can be obtained from twitter developer site. Twitter developed its own language amendment. Let's discussed the Twitter amendment [10]: -.
   a) Retweet convention of 'RT', that indicates the user is reposting or repeating.
   b) "#" used for hashtag to filter twitter messages i.e. called tweets according to topics or categories wise.
   c) "@user1" represents that reply of a message to a user whose user name is "user1" and so on.
   d) Length: Tweets are limited to 280 characters.
2. ***Data Pre-processing****:* In Data preprocessing we can remove irrelevant Twitter data. The steps in preprocessing of data are as follows-
   a) Case conversion: all words are converted either into lower case or upper case and Capital or smaller words in order to remove the difference between "Tweets" and "tweets" for further processing.
   b) Stop-words Removal: remove those words which have no meaning.
   c) Punctuation Removal: remove punctuation marks such as comma/colon which have no meaning.
   d) Spelling correction: incorrect words are automatically correct in this processing.
3. ***Feature Extraction***: After pre-processed of tweets words we need to extract features relevant for sentiment analysis. Following some feature including:
   a) Term frequency and Presence: it usually consists of n-grams of words and also their frequency counts.
   b) Parts of speech tagging: words in the text are tagged with their respective parts of speech in order to extract adjectives nouns verbs which ad meaning to the sentiment.

c) Opinion phrases and words: phrases or words that indicate opinion of the text.
d) Negation: presence of words like 'not', 'nor', 'neither' may reverse the sentiment of whole sentence. E.g.: "not good".
e) Twitter specific features: presence of emotions in tweets, positive or negative hashtags are all twitter specific features which add meaning to the sentiment [11].
f) Features exclusive to Tweets: the limited length of twitter which is 280 characters. The tweet specific features that have been exploited are listed below.
Emoticons: To represent happy and sad faces by using of emoticons. The presence of emoticons adds a weight to the feature which characterizes a positive or a negative sentiment.
Hashtags: these are features that allow you to group tweets related to a topic under a thread. The presence of hashtags for e.g. #android can express an orientation towards a positive or negative sentiment and hence the feature is exploited.
Punctuations: presence of marks, inverted commas (""), question marks(?), punctuations expressing ambiguity are all weighted to represent a feature. This feature represents a weight corresponding to the contribution of all punctuations in a review [5].

Tweet Mood Web Application: - [19] A Web application which performed real-time Sentiment analysis on Tweets of Twitter, that are matched particular matched keywords provided by users. Here discussed three ways of performing sentiment analysis: -

(i) **TweetScore: -** TweetScore calculates the polpularity of keywords of Twitter. This popularity number from 100 to -100. If more positive popularity score, suggest that the keyword is highly TweetScore and Popularity on twitter and other side if more negative popularity then lower TweetScore.
(ii) **TweetCompare: -** The popularity score compares the two or three different keywords on Twitter.
(iii) **TweetStats: -** TweetStats feature is for long term Sentiment Analysis. In this input number of popular keywords on Twitter on which operation runs after every hour in a day, calculates the popularity score for the tweets an hour time frame and store the results in database.

VI. MACHINE LEARNING ALGORITHMS FOR SENTIMENT CLASSIFICATION

A. Naïve Bayes classifier

[8] Naïve Bayes has better precision compared to other classifiers. It has lower accuracy and recall. The Naïve Bayes classifier is the simplest (as the name suggest) and most commonly used classifier. Other classifier has similar accuracy, precision and recall. There classifiers obtained an accuracy of 90% and other side Naïve Bayes has 89.5% accuracy Naive Bayes classifier works very well for text classification. for feature extraction bags of word are used in this model. Bayes Theorem a given feature:

P(label/features) =P(features/label) *P(label)/P(features)

P(label) describe prior probability a label or observed label. Given a feature, P(feature/label) is the prior probability that feature set of a given label.

B. Support Vector Machine (SVM)

[15] SVM finds the linear separators or hyperplane in the search for binary classification. There can be several hyperplanes that separate the classes, but the one that is chosen is the hyperplane in which each data points normal distance are large, so it proceed the maximum margin of Separation. Support vector machine is to detect the sentiments of Tweets [16].

C. Decision trees

Here, the training data space is represented in a hierarchical form in which shape on the attribute used to separation the data. The terms on attribute values are existence/non-appearance for single word or multiple. The separation of the data capacity done recursively until the minimum number of records are used [12]. Training and testing machine learning classifiers: After features are selected a machine learning classifier is chosen for sentiment analysis. Training data is used to train the classifier and its performance is measured using test data [8].

TABLE1

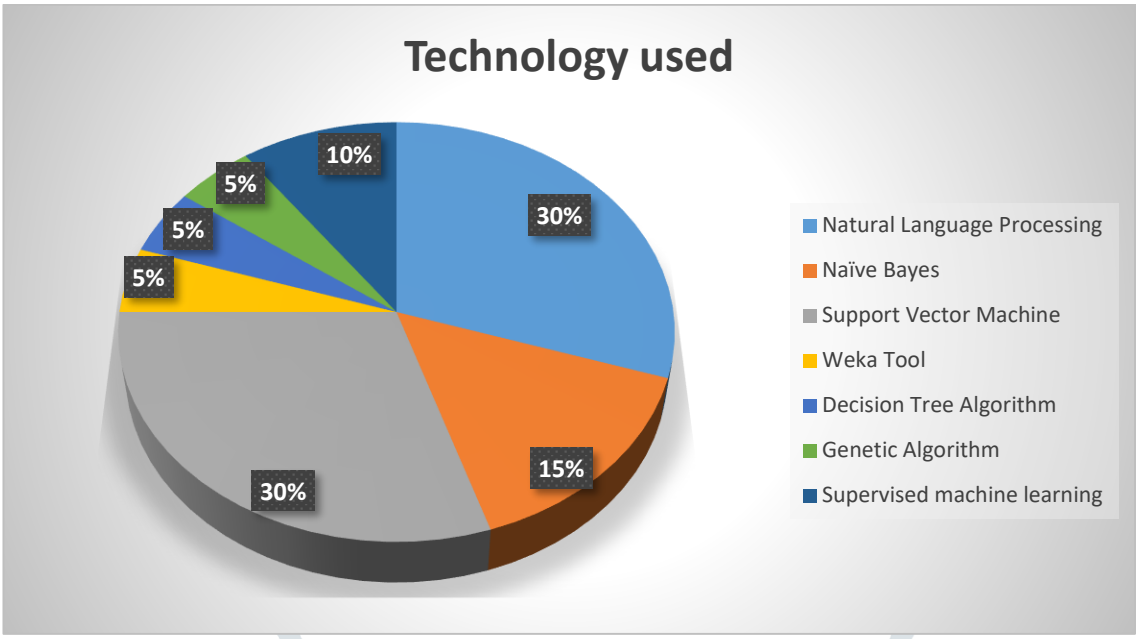
Overview of different papers shown in table 1: -

S.no	Paper title	Year of publication	Technique used	Highlights
1	An overview on opinion mining techniques	2018	Machine learning, Natural language processing	Sentiment or opinion classification techniques and mainly focus on sentiment types- regular type and comparative type
2	Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts	2018	Support vector machine, ANOVA	Subcategory median ranking of Analysed Facebook metrics by the sampled YouTubers.
3	Machine learning based Sentiment Analysis for Twitter Accounts	2018	Naïve Bayes	In this paper calculates sentiments from three analyzers that are SentiWord Net, TextBlob and W-WSD
4	Sentiment Analysis on Twitter using streaming API	2017	Natural language processing	Classification based on Naïve Bayes, time variant analytics and continuous leraning system. In this paper two limitations are present, first is Uni-Gram Naïve and second one is only used for English language
5	Machine learning & its application	2017	Support vector machine	Applications of Machine learning
6	A comprehensive Survey on Sentiment analysis	2017	Supervised Machine Learning	Opinion mining, sentiment analysis, social media
7	A survey of Sentiment analysis techniques	2017	Support vector machine	Classification techniques used for sentiment analysis
8	Sentiment analysis of product reviews: A Review	2017	Natural language processing	Machine learning algorithms and polarity-based sentiment classification
9	A Critical Review of Sentiment Analysis	2016	Natural language processing	Levels of classification and objectives of sentiment analysis
10	A brief review on Sentiment analysis	2016	Naïve Bayes classification	Different methods of sentiment analysis and different methods of sentiment levels

11	Exploring Sentiment Analysis on Twitter data	2015	Support vector machine	Sentiment analysis, SVM, lexicon, twitter
12	On Multi-Tier Sentiment analysis using supervised machine learning	2015	Supervised Machine Learning	Multi- tier classification Architecture in which data collection, data preprocessing, future extraction
13	The process of sentiment analysis: A study	2015	Natural language processing	Sentiment analysis, lexicon, social media
14	A Statistical & Evolutionary Approach to sentiment Analysis	2014	Genetic algorithms	Genetic algorithm used that explores a search space consisting of subsets of words in candidate paradigm words. Classify tweets via statistical method, where paradigm words are selected by genetic algorithm
15	Application of Machine Learning Techniques to Sentiment Analysis	2014	Support vector machine	Machine learning used NLP techniques that preprocessed the noisy data and the size of lexicon increases, this approach, more erroneous and time
16	Twitter sentiment analysis	2014	Natural language processing used	Twitter, sentiment analysis, opinion mining, social media, Natural Language Processing
17	Sentiment Analysis in twitter using Machine Learning	2013	Naïve Bayes	Symbolic techniques and sentiment classification techniques are mainly role in this research paper
18	Is the Sample Good Enough? Comparing data from Twitter's Streaming API with Twitter's fire hose	2013	SVM, Machine learning	Streaming API data, finding the no's of tweets matching, next statistical differences between two datasets
19	Twitter Sentiment Analysis	2012	Weka Machine Learning tool	Implementation and result discussion, in which results from objective/subjective classification and results from polarity classification (positive/negative)
20	Overview of use of Decision tree algorithms in Machine Learning	2011	Decision tree algorithm	Study of machine algorithms IDA and ID3

In table 1, discussed different papers in which various techniques of classification, analysis, algorithms and technologies are used. In above table we examine that mostly used natural language processing and support vector machine. SVM classifier, more popular and widely used for classifier. Naïve Bayes classification is a classifier that are used in above papers. NB classifier best and simple classifier type. Here, labels and feature are performed. In this table

describe each paper title, year of publish, technology used and highlights of papers. In pie chart analysis of technology used in this review paper. According to pie chart we analyzed that natural language processing technique 30% used in this and Support Vector Machine 30% also used in this paper. Naïve Bayes 15% used and supervised machine learning 10% used. And other remaining technologies are 5% used each in paper. This table describe the highlights of this paper. This able also used to analyzed this paper.



VII. CONCLUSION

In Sentiment Analysis using several types of Machine Learning Algorithms. In this paper researchers defined different algorithms accuracy, for example- Accuracy of Naïve Bayes 89.5% and others classifiers have accuracy 90%. In above table Natural Language Processing in Machine Learning for Sentiment Analysis focus on Regular and Comparative types. In this paper also used Streaming API for finding the no's of matching data in Social media site like Twitter. Machine learning used NLP for preprocessed of noisy data in Twitter. Some Other techniques of Machine learning like SVM, NLP, ME, NB etc. performed also their own duties. In this paper most usually, techniques are natural language processing and support vector machine. Other techniques are useful but few used in this review paper.

REFERENCES

[1] Balaji, Penubaka, D. Haritha, and O. Nagaraju. "An Overview on Opinion Mining Techniques and Sentiment Analysis." *International Journal of Pure and Applied Mathematics* 118, no. 19 (2018): 61-69.

[2] Poecze, Flora, Claus Ebster, and Christine Strauss. "Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts." *Procedia computer science* 130, no. C (2018): 660-666.

[3] Hasan, Ali, Sana Moin, Ahmad Karim, and Shahaboddin Shamshirband. "Machine learning-based sentiment analysis for twitter accounts." *Mathematical and Computational Applications* 23, no. 1 (2018): 11

[4] Trupthi, M., Suresh Pabboju, and G. Narasimha. "Sentiment analysis on twitter using streaming API." In *2017 IEEE 7th International Advance Computing Conference (IACC)*, pp. 915-919. IEEE, 2017.

[5] Angra, Sheena, and Sachin Ahuja. "Machine learning and its applications: A review." In *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, pp. 57-60. IEEE, 2017.

[6] Rajalakshmi, S., S. Asha, and N. Pazhaniraja. "A comprehensive survey on sentiment analysis." In *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, pp. 1-5. IEEE, 2017.

[7] Kaur, Harpreet, and Veenu Mangat. "A survey of sentiment analysis techniques." In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, pp. 921-925. IEEE, 2017.

[8] Shivaprasad, T. K., and Jyothi Shetty. "Sentiment analysis of product reviews: A review." In *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 298-301. IEEE, 2017.

[9] Chopra, Fatehjeet Kaur, and Rekha Bhatia. "A Critical Review of Sentiment Analysis." *International Journal of Computer Applications* 149, no. 10 (2016).

[10] Raghuvanshi, Neha, and J. M. Patil. "A brief review on sentiment analysis." In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pp. 2827-2831. IEEE, 2016.

[11] Venugopalan, Manju, and Deepa Gupta. "Exploring sentiment analysis on twitter data." In *2015 Eighth International Conference on Contemporary Computing (IC3)*, pp. 241-247. IEEE, 2015.

[12] Moh, Melody, Abhiteja Gajjala, Siva Charan Reddy Gangireddy, and Teng-Sheng Moh. "On multi-tier sentiment analysis using supervised machine learning." In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 1, pp. 341-344. IEEE, 2015.

[13] Godsay, Manasee. "The process of sentiment analysis: a study." *International Journal of Computer Applications* 126, no. 7 (2015).

[14] Carvalho, Jonnathan, Adriana Prado, and Alexandre Plastino. "A statistical and evolutionary approach to sentiment analysis." In *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02*, pp. 110-117. IEEE Computer Society, 2014.

[15] Jain, Anuja P., and Padma Dandannavar. "Application of machine learning techniques to sentiment analysis." In *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, pp. 628-632. IEEE, 2016.

[16] Sarlan, Aliza, Chayanit Nadam, and Shuib Basri. "Twitter sentiment analysis." In *Proceedings of the 6th International Conference on Information Technology and Multimedia*, pp. 212-216. IEEE, 2014.

[17] Neethu, M. S., and R. Rajasree. "Sentiment analysis in twitter using machine learning techniques." In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pp. 1-5. IEEE, 2013.

[18] Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and Kathleen M. Carley. "Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose." In *Seventh international AAAI conference on weblogs and social media*. 2013.

[19] Baqapuri, Afroze Ibrahim. "Twitter sentiment analysis." *arXiv preprint arXiv:1509.04219* (2015).

[20] Navada, Arundhati, Aamir Nizam Ansari, Siddharth Patil, and Balwant A. Sonkamble. "Overview of use of decision tree algorithms in machine learning." In *2011 IEEE control and system graduate research colloquium*, pp. 37-42. IEEE, 2011.