Survey On Text Categorization Using Sentiment Analysis

Chaitanya Bhagat, Dr. Deepak Mane

Abstract— Twitter is a blog website online on internet which offers the platform to humans to experience and talk their perspectives about troubles, occurrences, merchandise and exclusive mind. Sentiment Analysis is an open-ended subject of research in the text mining area. Using several systems gaining knowledge of algorithms for evaluation of sentiments from exceptional tweets having a most of one hundred forty phrases consistent with a tweet and proposes a studies technique for improvisation of class. This survey paper tries to provide an entire evaluation of the modern replace in this discipline. The most important goal of this survey is to give a complete picture of ways Machine studying strategies are used in Sentiment Analysis to get better effects in short details. Also, we will have a look at basic emotion's classification into ternary lessons i.e. Fantastic, negative, neutral the usage of exclusive device learning algorithm and type into their subclasses i.e. Love, happiness, fun, neutral, hate, sadness, and anger the use of equal system getting to know algorithms.

Index Terms— K-Nearest Neighbors, Machine Learning, Naïve Bayes, sentiment analysis, Twitter.

1 INTRODUCTION

TWITTER kept some traits that make it an exceedingly exciting state of affairs of information mining. In its primary shape, Twitter is a micro-blogging carrier that allows customers to put up quick textual content updates, with the only belongings of not allowing more than a hundred and forty characters in a single text message. This quandary became out to be a very appealing asset because it allows posting short, even contemporary-time, updates regarding one's activities and facilitates sharing and forwarding repute messages, in addition to replying to them fast. It has an advantageous effect or a negative one, whether the news spread is real or false, and so on. These surroundings provide a totally intense, source of facts to mine. Anyway, due to the constraint as a way as characters (i.E.One hundred forty characters in line with a tweet), mining such information presents lower performances than that after mining longer texts. In addition, partition into multiple classes stays a tough challenge: binary class of textual content usually relies on the sentiment polarity of its additives (either positive or negative) while, when positive and negative training is divided into subclasses, the accuracy tends to lower remarkably. In this new approach, we recommend an approach that is predicated on writing patterns, and unique unigrams to classify tweets into eight unique subclasses and describe how the proposed approach provides accurate performances. The objective of this machine is to analyses the distinctive tweet feelings in line with a different category. Calculate the most lively category on twitter like Sport, Entertainment, and Politics. Calculate these days young generation busy with Twitter and additionally Categorizations of the tweet in keeping with unique user tweet multiclass sentiment evaluation techniques.

2 RELATED WORK

M. Bouazizi and T. Ohtsuki [1] proposed a novel method to classify sentiment into binary and ternary training which might be predominant classes. And additionally, classify text accrued from Twitter into seven subclasses. For this, they created the SENTA tool to the user to select functions and then run type into seven different classes (fun, happiness, love, neutral, sadness, anger, hate). For tweet emotion type classification, they used Random Forest classifier on textual Twitter information and attain accuracy 70.1% on ternary classes and 60.2% on seven training. S. M. Mohammad and S. Kiritchenko [2] introduced hashtag emotion lexicon technique are used to generate a big lexicon of the word–emotion from this emotion-categorized tweet corpus and then further classification on guide labels of emotions in tweets. Although, due to the restriction as some distance as characters in this article, show that feeling word hashtags are high-quality guide marks of feelings in tweets. In those essay's information set is used and hashtag emotion lexicon, the SVM approach is used. An approach to provide an expansive vocabulary of the phrase– feeling relationship from this sense marked tweet corpus. This is the primary vocabulary with the real esteemed phrase– feeling affiliation scores. Start with investigations for six vital emotions and demonstrate that the hashtag causes are reliable and coordinate with the comments of organized judges. The device suggests how the extricated tweet corpus and word– feeling affiliations may be utilized to enhance feeling grouping precision in a change no tweet domain. Personality is probably related to any of the several emotions and due to the fact that our hashtag method scales effects to countless, increase our corpus through accumulating tweets with hashtags regarding 585 pleasant feelings. B. Plank and D. Hovy [3] delivered the usage of social media as an aid for huge-scale, open vocabulary personality detection like information, such as gender, number of fans, statuses, or listing club, upload treasured information. Brain technological know-how study proposes that certain identity characteristics relate to semantic conduct. This courting may be accurately tested with measurable feature dialect managing methods. Expectation precision, for the maximum component, enhances with larger facts exams, which moreover takes into consideration greater lexical highlights. Most existing work on identification expectation, anyhow, facilities on little examples and close vocabulary examinations. The two variables constrain the all-inclusive assertion additionally, the

measurable intensity of the effects. In this, gender-controlled dataset and Natural Language Processing method are used to investigating the usage of social media as an asset for expansive scale, open vocabulary identification vicinity. Dissect which highlights are prescient of which identification qualities and present a novel corpus of one. 2M English tweets commented on with Myers-Briggs identity compose and sexual orientation. their trials demonstrate that online networking facts can provide adequate semantic evidence to dependably predict 2 of 4 identity measurements. Ankur Goel et al. [4] states that tweets may be labelled using SentiWordNet together with Naive Bayes for the category into unique training based totally on their relevance with the subject searched. Twitter is a miniaturized scale blogging website online which gives a stage for people to percentage and specific their views approximately topics, happenings, items, and distinctive administrations. Tweets may be arranged into various lessons dependent on their importance with the theme sought. Different Machine Learning calculations are as of now utilized in order of tweets into wonderful and terrible lessons dependent on their estimations, for instance, Baseline, Naive Bayes Classifier, Support Vector Machine and so on. Twitter includes the usage of Naive Bayes using sentiment140 getting ready facts utilizing twitter database and recommend a method to make strides characterization. This system includes a big movie assessment dataset. In this actual-time sentiment analysis category approach is used. Utilization of SentiWordNet alongside Naive Bayes can beautify exactness of association of tweets, by way of giving energy, cynicism and objectivity rating of phrases present in tweets. Aparna Garimella and Rada Mihalcea [5] brought gender difference on social media. The surface-degree text category approaches to gender discrimination and attempted to benefit insights into the differences between women and men with the aid of using semantic methods that can point to salient phrase classes or differences in concept usage. Data is accumulated from the BlogSpot take any other take a look at the problem of gender discrimination and attempt to flow beyond the typical surface-degree textual content class approach, through figuring out semantic and psycholinguistic word lessons that mirror systematic variations between males and females and finding differences between genders inside the approaches they use the equal words. The machine describes numerous experiments and reports consequences on a big series of blogs authored by way of males and females. Believe those distinctions at a deeper semantic degree. D. Bamman and N. A. Smith [6] to observe of sarcasm and other speech act on social media web sites with complex audiences this contextualized sarcasm detection on twitter. Most computational tactics to sarcasm detection, however, deal with it as a merely linguistic matter, the usage of records which includes lexical cues and their corresponding sentiment. Sarcasm calls for a few mutual getting to know among speakers and amassing of people; it's miles a notably applicable surprise. Most computational approaches to cope with sarcasm area, anyways, regard it as an easy semantic problem, making use of facts, for example, lexical prompts and there pertaining to supposition as prescient functions. They display that via including additional etymological facts from the unique state of affairs of an articulation on Twitter, as an example, houses of the writer, the group of onlookers and the set off open case they can accomplish profits in precision contrasted within reality phonetic highlights in the discovery of this elaborate wonder, while moreover revealing perception on highlights of relational collaboration that empower sarcasm in dialogue. P. Praveen et al. [7] states approximately evaluations at the film. It's very complicated to move across all of the views, because the wide variety of opinions frequently posted for a movie in various social network websites, by using sentiment evaluation it is feasible to divide overall reviews into superb, bad and neutral opinions. The proposed system includes system learning strategies for the sentiment analysis, emotions of the humans with the assist of the R language, it could be used for predicting the repute of the strolling movies. In the sentiment analysis every word in the sentence categorized into three types including fine, terrible and neutral, this technique implemented using textual content pre-processing strategies and naïve byes approach for sentiment class. Apoorv Agarwal et al. [8] shown consequences for estimation research on Twitter. The device displayed a radical arrangement of examinations for each these assignments on bodily commented on information that is an arbitrary instance of a movement of tweets. User examined two forms of fashions: tree-based version and feature-based totally fashion those beats the unigram pattern. For thing-based total methodology, they highlight research which uncovers that the maximum substantial highlights are those that consolidate the earlier extremity of phrases and their grammatical forms labels. The new gadget likely presumes that slant research for Twitter statistics is excessively no longer pretty the same as opinion investigation for different sorts. M. Bouazizi and T. Ohtsuki [9] states that fine in elegance methodologies of opinion examination present several Weaknesses whilst arranging tweets. Multi-elegance notion exam has tested to be a tough project. This is chiefly for the straightforward motive that a tweet commonly does not incorporate a solitary evaluation, but a huge quantity. In this, they advocate an example-based technique for evaluation size on Twitter. By evaluation, they find the existing estimations interior a tweet and the invention of the burden of these estimations. Initially, tweets are grouped into nice, bad, or nonpartisan. This technique achieves an exactness of 81%. They characterize 2 measurements to degree the accuracy of feeling discovery and exhibit that supposition measurement can be an extra crucial undertaking than the commonplace multi-class characterization. Paolo Ferragina et al. [10] provided the Hashtag-Entity Graph and appropriate calculations to strongly cope with IR problems deliberate on Twitter hashtags: relatedness and characterization. They tried our calculations over known and new datasets, drawn from Twitter, whose look at is to two requests of greatness bigger than the present day ones. These huge datasets were discharged to people in popular, collectively with the graph we constructed. They contend that the HE-chart gives a concise but floor-breaking portrayal for tweets, pleasantly mixing semantic relatedness among substances what is extra, co-occasion data among hashtags and people substances. They function two extensive residences of our calculations: (i) they may be language self-sufficient (as much as a semantic annotator is obtainable for the goal language) and (ii) can be applied in an online framework, in which the clients conceivably talk approximately new materials. Wei Gao and Fabrizio Sebastiani [11] proposed Opinion class has turned into a commonplace empowering innovation within the Twittersphere. Characterizing

tweets as according to the belief they pass on closer to a given substance has numerous packages in political principle, sociology, promote it observe, and numerous others. Meanwhile, right here, the battle that most beyond investigations managing tweet assessment association utilize a less than excellent methodology. The motive is that the last goal of maximum such examinations aren't assessing the classmark (e.G., Positive, Negative, or on the other hand Neutral) of man or woman tweets, but assessing the relative recurrence of the superb lessons within the dataset. The last undertaking is called dimension, and past due research has convincingly appeared that it needs to be handled as its very very own errand, using studying calculations and assessment estimates specific from the ones applied for characterization.

Table 1: Summarization of several methods for sentiment analysis

Year	Paper Title	Summary	Method, Dataset and Accuracy Achieved
2017	A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter [1].	In this technique, Random Forest set of rules used for tweet classification into binary and multiclass classification.	**Method:** Random Forest **Dataset:** Twitter data set **Accuracy:** 81.32% for two class, 70.1% for three class and 60.2% for seven classes.
2017	Sentiment Analysis on Twitter Data using KNN and SVM [13].	Using KNN and SVM with a couple of features from tweet sentiment is detected.	**Method:** Support Vector Machine and K-Nearest Neighbor **Dataset:** Twitter dataset **Accuracy:** 84.32% with KNN And 77.99% with SVM
2016	Real Time sentiment Analysis of Tweets Using Naïve Bayes [4].	Tweets can be categorised the usage of sentiWordnet alongside Naive Bayes into terrific lessons based totally on their relevance with the topic. searched.	**Method:** Naïve Bayes **Dataset:** Large Movie Review Dataset(LMRD) **Accuracy:** 58.40%
2015	Using Hashtags to Capture Fine Emotion Categories from Tweets [2].	Using Hashtag phrases of tweets, the massive lexicon of word-emotion is generated. And with the aid of the usage of SVM on Facebook records set to perceive emotion is both good or bad.	**Method:** Support Vector Machine **Dataset:** Facebook dataset **Accuracy:** 50%
2015	Contextualized Sarcasm Detection on Twitter [6].	Sarcasm detected from tweet by extracting features and then binary linear regression is applied.	**Method:** Binary Logistic Regression **Dataset:** Twitter dataset **Accuracy:** 85.1%
2013	Sentiment analysis of Hollywood movies on Twitter [12].	To examine the sentiments expressed on Hollywood movies on Twitter in order that customer's evaluations, behavior and choices are extracted, analyzed and used to understand the marketplace behavior for better consumer experience.	**Method:** Naïve Bayes **Dataset:** Twitter dataset **Accuracy:** 79%
2018	Deep Convolution Neural Networks for Twitter Sentiment Analysis [17].	In this technique, they introduce the word embedding technique. Words embedding are joined with n-grams functions and phrase sentiment polarity score abilities to shape the sentiment function set of tweets. The function set is blended to a deep convolution neural community for sentiment category.	**Method:** Convolution Neural Network **Dataset:** 1) The Stanford Twitter Sentiment Test data set 2) The SE2014 dataset 3) The Stanford Twitter Sentiment Gold data set 4) The Sentiment Evaluation dataset 5) The Sentiment Strength Twitter dataset **Accuracy:** 87.62%

3 FEATURE SELECTION IN SENTIMENT CLASSIFICATION

Sentiment Analysis purpose is sentiment classification problem. The first step within the Sentiment Classification problem is to obtain and decide text capabilities.

3.1 Terms presence and frequency

These features are specific words or word n-grams and their frequency counts. It both offers the words binary weighting (0 if the phrase resembles, or one if otherwise) or uses time period frequency weights to expose the relevant importance of features.

3.2 Parts of speech (POS)

Obtaining adjectives, as they are valuable symptoms of evaluations.

3.3 Opinion words and phrases

These are words commonly used to formulate opinions which include correct or bad, like or hate. On every other hand, a few phrases constitute opinions without utilizing opinion phrases. For example: cost me an arm and a leg.

Negations

The presence of terrible words may ruin the opinion familiarization like no longer exact is equivalent to horrific.

4 SENTIMENT CLASSIFICATION TECHNIQUES

Sentiment Classification strategies arranged into machine Learning, lexicon-based approach and hybrid technique. The Machine Learning Approach employs famous Machine Learning algorithms and uses semantic functions. The Lexicon-based Approach relies on a sentiment lexicon, a set of identified and precompiled sentiment phrases. It is split right into a dictionary-primarily based method and corpus-based totally method. The hybrid Approach merges different methods and may get better results in sentiment analysis. The text classification methods using the Machine Learning techniques can be roughly classified into supervised and unsupervised learning methods. The supervised techniques make usage of a huge number of labelled training records. The unsupervised techniques are employed when it is hard to find these labelled training records. The lexicon-based approach depends on getting the opinion lexicon which is utilized to analyze the text. There are two methods in this approach. The dictionary-based approach which depends on obtaining opinion roots words, and then examines the dictionary of their synonyms and antonyms. The corpus-based approach starts with a roots list of opinion words and then obtains another opinion words in a huge corpus to support in obtaining opinion words with context-specific orientations. This could be achieved by applying statistical or semantic approaches.

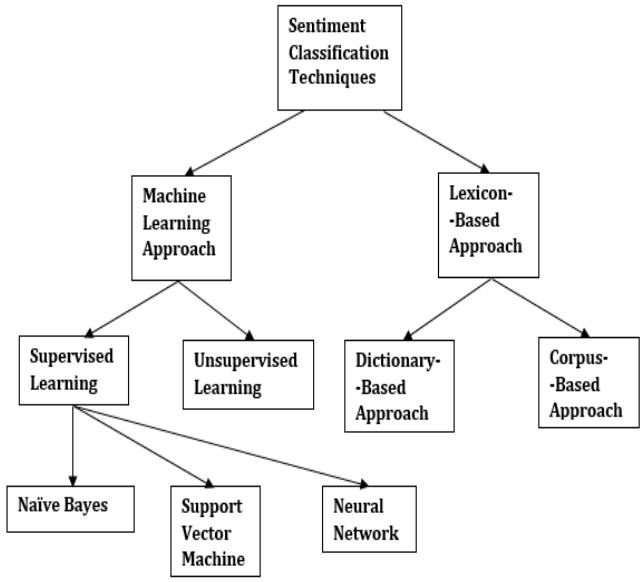

Figure 1. Tree diagram of sentiment classification techniques

4.1 Machine learning approach

Sentiment Analysis as a universal text classification problem that enforces the use of syntactic plus linguistic features. Text Classification Problem Definition: We have a set of training documents Doc = {Y1, Y2, …, Yn} where each document is labelled to a class. The classification model is appropriate to the features in the underlying record to classify according to the class label. Next for a given instance of an undiscovered class, the model is utilized to predict a class label for it. The hard classification problem is when only one label is selected to an instance. The soft classification problem is when a probabilistic value of labels is selected to an instance.

4.1.1 Supervised learning

The supervised strategies rely upon the presence of labelled training data. There are many forms of supervised classifiers some of the most frequently used classifiers in Sentiment Analysis are Naive Bayes, Support Vector machine, and Random Forest.

4.1.1.1 Naïve Bayes

Naive Bayes is a simple and easy but powerful category set of rules. It is primarily used for file-stage classification. Naive Bayes classifiers are computationally rapid whilst taking selections. Naïve Bayes Classifier is well used in real problems together with electronic mail Spam detection, Sentiment analysis, and Sexual content detection. It is fairly advocated when we have a lack of resources in phrases of reminiscence and a Central processing unit (CPU). It desires low processing memory and less execution time. Naïve Bayes type model estimates the posterior possibility of a category, based on the distribution of the words in the report. The model works with the Bag Of Words function extraction which ignores the position of the phrase inside the record [4].

4.1.1.2 Support Vector Machine

Support Vector Machine is a discriminative classifier considered because of the satisfactory text classification approach. It is a statistical classification technique SVM maps input (actual-valued) feature vectors right into a higher-dimensional feature area via a few nonlinear mapping. SVMs are developed on the precept of structural hazard minimization. The structural hazard minimization seeks to discover a hypothesis (h) for which one can locate the bottom probability of blunders whereas the conventional mastering strategies for pattern reputation are primarily based at the minimization of the empirical chance, that's a try and optimize the performance of the gaining knowledge of set. Computing the hyper aircraft to split the statistics points i.e. Education and Support Vector Machine (SVM) results in quadratic optimization trouble. Support Vector Machine can study a bigger set of patterns and capable of scale higher, because of class complexity it does not depend on the dimensionality of the function space. Support Vector Machine (SVM) has the capacity to replace the schooling styles dynamically every time there may be a brand new sample throughout category [2].

4.1.1.3 Neural Network

Neural Networks are a class of fashions within the widespread gadget gaining knowledge of literature. Neural networks are a particular set of algorithms which have transformed machine learning. They are inspired with the aid of organic neural networks and the popular so-called deep neural networks have shown to work quite properly. Neural Networks are themselves standard feature processes, that is why they can be hired to almost any device learning trouble regarding studying a complex mapping from the entrance to the output area. Neural network-based strategies have executed brilliant improvement in an expansion of herbal language processing obligations [15].

4.1.2 Unsupervised learning

It does no longer consist of a category and that they do not provide with the correct objectives in any respect and as a result depend on clustering.

4.2 LEXICON-BASED APPROACH

Opinion phrases are used in lots of sentiment class assignments. Positive opinion words are applied to reveal some favored activities, even as terrible opinion phrases are hired to specify a few undesired events. There also are opinion phrases and idioms which together are known as opinion lexicon. There are 3 crucial methods on the way to deliver collectively or accumulate the opinion phrase list. The manual method may be very time arduous and it isn't used alone. It is typically joined with the alternative computerized procedures as a very last check to keep away from the errors that resulted from automated techniques.

4.2.1 Dictionary-based approach

A small set of opinion words is obtained manually with known orientations. Then, this set is increased with the aid of searching inside the well-known corpora WordNet or word list for their synonyms and antonyms. The newly obtained words are mixed to the basic list then the subsequent iteration starts off evolved. The iterative system stops whilst no new word are detected. After the process is finished, the manual examination can be taken out to eliminate or correct errors. The dictionary-primarily based approach has the principle downside that is the weak spot to locate opinion words with discipline and context-specific orientations.

4.2.2 Corpus-based approach

The Corpus-based method serves to solve the hassle of acquiring opinion words with context-specific orientations. Its methods depend upon syntactic patterns or styles that arise typically along with a root listing of opinion words to discover other opinion words in a big corpus.

5 EVALUATION OF SENTIMENT ANALYSIS CLASSIFICATION

Using Accuracy, Precision and Recall we can evolution of sentiment analysis of twitter**.**

5.1 Accuracy

Accuracy refers to the overall correctness of classification. It measures the ratio of correctly classified instances over the total number of instances.

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Where TP is True Positive, TN is True Negative, FP is false positive, and FN is False Negative. Accuracy is the ratio of Total positive and Total negative to Total.

5.2 Precision

Refers to the fraction of the tweets correctly classified, for a given sentiment, over the total number of tweets classified as belonging to that sentiment.

Precision = TP/(TP+FP)

5.3 Recall

Refers to the fraction of tweets correctly classified, for a given sentiment, over the total number of tweets actually belonging to that sentiment.

Recall = TP/(TP+FN)

6. CHALLENGES IN SENTIMENT ANALYSIS

6.1 Language Problem

In Opinion mining, the English language is pretty well utilized because of its resource availability means lexicons, dictionaries and corpora but researchers become interested in using Opinion mining with a language other than English (Arabic, Chinese, German, etc). Hence, researchers encounter a challenge for creating resources i.e. lexicons, dictionaries and corpora for these languages.

6.2 Natural Language Processing (NLP)

Using NLP in the Opinion Mining process requires more improvements because it attracts researches. And NLP offers better Opinion Mining results and provides good language perception. There is a need to spend more attention to the

research of domain-dependent opinion mining or context- based opinion mining because domain-specific Opinion Mining provides good result than domain independent corpus. And domain specific Opinion Mining is complex or more difficult to develop.

6.2 Fake Opinion

It is also termed Fake review and refers to bogus or fake reviews which misguide the readers or consumers by presenting them fraudulent negative or positive opinions associated with an object and in order to lower the status of an object. These spams make sentiment opinion ineffective in many application areas.

7. CONCLUSION

In sentiment analysis mostly machine learning approach are used we can see in [1], [2], [4], [12] and [13]. Machine learning approach having many classification algorithm and each algorithm have different computational power than each other. We can see dataset mostly consists of user comments on social media i.e. Twitter dataset, Facebook dataset or users reviews on movie which is written on any web site. Sentiment can be classified into main categories (Positive, Negative and Neutral) as well as sub categories (Fun, Happiness, Love, Neutral, Hate, Sad, Anger). Random Forest, Support Vector Machine, Naïve Bayes and Logistic Regression these are machine learning algorithms which gives highest sentiment classification accuracy is 81.3 %, 77.99%, 79% and 85.1% respectively. This survey paper presented an overview of the modern updates in Sentiment Analysis algorithms and applications. After analyzing these, it is clear that the enhancements of Sentiment Classification are still an open field for research. Naïve Bayes and Support Vector Machines are the most regularly used Machine Learning algorithms for determining Sentiment Classification problem. They are supposed a reference model where many proposed algorithms are compared to. Interest in languages apart from English in this field is increasing as there may be nonetheless a shortage of resources and researches concerning those languages. Information from micro-blogs, blogs, and boards in addition to a news source, is extensively utilized in Sentiment Analysis recently. This media information plays a vast role in showing people's sentiments, or opinions about a particular topic or product. Using social network sites and micro-blogging sites as a root of data still need broader analysis. There are some benchmark data sets particularly in reviews like IMDB which are utilized for algorithms evaluation.

REFERENCES

[1] M. Bouazizi and T. Ohtsuki, "A Pattern-Based Approach for Multi-class Sentiment Analysis in Twitter," in Proc. IEEE ACCESS, pp. 20617-20639, 2017.

[2] S. M. Mohammad and S. Kiritchenko, "Using Hashtags to Capture Fine Emotion Categories from Tweets," in Computational Intelligence, vol. 31, no. 2, 2015, pp. 301–326.

[3] B. Plank and D. Hovy, "Personality Traits on Twitter or How to Get 1,500 Personality Tests in a Week," in Proc. of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2015, pp. 92–98.

[4] Ankur Goel, Jyoti Gautam, Sitesh Kumar, "Real Time Sentiment Analysis of Tweets Using Naive Bayes," in 2nd International Conference on Next Generation Computing Technologies, Dehradun, 2016.

[5] Aparna Garimella and Rada Mihalcea, "Zooming in on Gender Differences in Social Media," in Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media, 2016.

[6] D. Bamman and N. A. Smith, "Contextualized Sarcasm Detection on Twitter," in Proc. of the 9th Int. AAAI Conf. on Web and Social Media Citeseer, 2015, pp. 574–577.

[7] P. Praveen, P. Sudheer and K.Sudheer Kumar, "Public Sentiment Analysis On Movie Reviews," in Computer Science and Engineering department, 2018.

[8] Apoorv Agarwal, Boyi Xie,Ilia Vovsha, Owen Rambow, Rebecca Passonneau, "Sentiment Analysis of Twitter Data," in Department of Computer Science Columbia University New York, 2016.

[9] M. Bouazizi and T. Ohtsuki, ''Sentiment analysis in Twitter: From classification to quantification of sentiments within tweets,'' in Proc. IEEE GLOBECOM, 2016, pp. 1–6.

[10] Paolo Ferragina, Francesco Piccinno, Roberto Santoro, "On Analyzing Hashtags in Twitter," in Dipartiment of di Informatica University of Pisa Proceedings of the Ninth International AAAI Conference on Web and Social Media, 2016.

[11] Wei Gao and Fabrizio Sebastiani, "Tweet Sentiment: From Classification to Quantification," in Qatar Computing Research Institute Hamad bin Khalifa University PO Box 5825, Doha, Qatar IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2015.

[12] U. R. Hadeghatta, "Sentiment Analysis of Hollywood movies on Twitter," in Proc. IEEE/ACM ASONAM, 2013, pp. 1401-1404.

[13] Mohammad Rezwanual Huq, Ahmad Ali, Anika Rahman, "Sentiment Analysis on Twitter Data using KNN and SVM," in Proc. IJACSA, pp 2381-2388, 2017.

[14] Raheesa Safrin, K.R. Sharmila, T. S. Shri, Subangi, E. A. Vimal, "Fraud Detection of Facebook business page based on sentiment analysis," in IRJET, pp. 2381-2388.

[15] Penqfei Liu, Xipenq Qiu, Xuanjing Huanq, "Research on Short Text Classification Algorithm Based on Neural Network," in 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, pp. 1726-1729, 2018.

[16] Radhi Desai, "Sentiment Analysis of Twitter Data: A Survey," in International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2018, pp. 464-470.

[17] Zhao Jianqiang, Gui Xiaolin, Zhang Xuejun, "Deep Convolution Neural Networks for Twitter Sentiment Analysis," in IEEE ACCESS, 2018, pp. 23253-23260.

[18] Bala Durga Dharmavarapu, Jayanag Bayana, "Sarcasm Detection in Twitter using Sentiment Analysis," in International Journal of Recent Technology and Engineering, ISSN: 2277-3878, Volume-8, Issue-1S4, 2019,

pp. 642-644.
[19] Daneena Deeksha Dsouza, Deepika, Divya P. Nayak, Elveera Jenisha Machado, Adesh N. D "Sentimental Analysis of Student Feedback using Machine Learning Techniques," in International Journal of Recent Technology and Engineering, ISSN: 2277-3878, Volume-8, Issue-1S4, 2019, pp. 986-991.
[20] Siva Parvathi, V, Yjn Lakshmi, "A Survey on Sentiment Analysis," in International Journal for Innovative Engineering & Management Research, Vol. 7, No. 12, 2018, pp. 289-296.