Research Article

Emotion And Sentiment Analysis From Twitter Text

[1]amulya Chenna, [2] B.Srinivas, [3]s.Nagaraju

[1]M.Tech student, Dept. of CSE, Kakatiya Institute of Technology& Science, Warangal, TS
[2]Assistant Professor, Dept. of CSE, Kakatiya Institute of Technology& Science, Warangal, TS
[3]Assistant Professor, Dept. of CSE, Kakatiya Institute of Technology& Science, Warangal, TS
[1]amulya.chenna508@gmail.com, [2]bs.cse@kitsw.ac.in, [3]sn.cse@kitsw.ac.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 May 2021

Abstract: Social media like Twitter and Facebook are full of activities, emotions, and reviews from customers worldwide. Emotions are recognized improved through their adaptive use in maintaining basic services and lifestyles. Each emotion has individual characteristics: Anger, Surprise, Joy, Fear, disgust, sadness, etc. Each emotion also has characteristics in community with other emotions. We can distribute our moments, expressions, ideas, and mental state, national and international difficulties within the textual content, images, messages, and audio and video posts. Although we have different conversation types, text transmission is one of the most popular conversation techniques on social media. This paper aims to describe and understand the sentiment and emotion that humans have discovered from the text of their social media posts and apply it to generate advice. This paper collected tweet information and replies on a few special points and created a dataset with text, emotion, sentiment data, etc. We used 500 tweets data set to obtain feelings and emotions in Tweets and their re-tweets, and we covered user impact scores based on various people-based measures and tweets. The machine learning algorithm plays a vital role in sentiment analysis. In this paper, a machine learning-based Naïve Bayes classifier is used to perform emotion-relevant text classification like Anger, surprise, joy, fear, sadness, etc. The experimental results show that the proposed algorithm provides better accuracy compared to previous algorithms of KNN and DT classifiers.

Keywords: Emotion, Sentiment, Text, Emotion models, Emotion detection, Sentiment detection, Emotion analysis, Sentiment analysis.

1. Introduction

In the past few years, there has been a vast upward momentum in the online shopping market. The online shopping market has transformed from a general insurance market to a multi-billion dollar industry. In this world, humans buy a product either in-store or online. In the new world, most people choose to buy a product online because of its many blessings. Buying things and goods online has become a new way of shopping for thousands and thousands of people worldwide. There is a massive growth in the diversity of people buying the product online. This huge growth in online shopping is because there are lower prices, more savings options, a giant competition, and more product shape; it saves cases and is delivered in its wake.

Sentiment analysis is also known as 'opinion mining' or 'artificial intelligence of emotions' and refers to the use of natural language processing (NLP), text exploration, computational linguistics, and biometrics to systematically recognize, extract, examine and monitor emotions. States and subjective facts. Sentiment analysis is generally concerned with sound in consumables; For example, surveys and reviews on the web and social networks based on the entire web. As a general rule, sentiment analysis attempts to determine the behavior of the speaker, author of articles, or various topics in topic sentences through intense emotional or emotional responses to a file, statement, or occasion. This behavior is likely to be a judgment or evaluation, full of emotion (in other words, the emotional attitude of the writer or speaker) or the expectation of enthusiastic responses (in other sentences, the influence of the writer or buyer use). Today, there is a large number of customer polls or suggestions on the web on all topics, and audits may also contain surveys about objects, including customers, movie bugs, etc. View on the web. Large numbers of surveys are available for single organisms, making this difficult for clients because they need to check all of them for you to choose. Thus, extracting these records, distinguishing and organizing customer opinions is an important task. Sentiment prospecting is a task that leverages natural language processing and fact extraction (IE) procedures to investigate in depth a variety of files in an effort to synthesize feedback with the help of unique authors [1].

Today, the Internet is the most popular medium for conversation, and it is rich in feelings. Feelings can be extracted from textual input by reading punctuation marks, emotional keywords, grammatical form, and semantic facts with natural language processing strategies. There is a great deal of text data on the Internet. It's exciting to extract feelings for dreams as unique as those of a business. For example, in luxury goods, emotional factors such as brand, strength, and prestige for purchasing decisions are more important than rational elements that include technical, beneficial, or definitions. In this example, the customer is happy to purchase a product despite the increased fees. Emotional marketing aims to stimulate emotions in the business owner to relate it to the brand and thus enhance the promotion of the product/supplier. Nowadays, it is not the product to be purchased, as there is a lot of preference for each category. Still, interest is the buyer's relationship with the slogan and feelings that the product communicates with [2].

Twitter is a large and rapidly developing small site that runs a blogging social networking website where people post their reviews in a quick and simple expressive way. It is common for merchants selling products on the web to ask their customers to verify products. On Twitter, several customer reviews are working on various goods. Electronic good is a popular area where a lot of customer reviews pop up. Among the electronic products, cell phones are a major area where customers enjoy a great hobby. It makes it difficult for the buyer to inspect and decide whether to purchase the product.

2. Review of literature

Similar to sentiment, emotions can be analyzed computationally. However, the goal of emotion analysis is to recognize the emotion, not the sentiment, which makes it an additional difficult challenge, as the differences between sentiments are more understated than those between the massive and the terrible. Although the evaluation of feelings and emotions are separate duties, our assessment of the literature shows that the use of either time period is not always consistent. There are instances when researchers analyze the negative aspects and the best quality of the text; however, they refer to their analysis as an emotion analysis. Likewise, there are cases where researchers investigate a set of subjective feelings that include feelings, but call them sentiment analysis.

Sheresh Zahoor et al. [2020] Sentiment analysis, also known as opinion-eliciting or sentiment extraction, is a sentiment within text records. This technique has been used widely over time to identify feelings and feelings within specific textual facts. Twitter is a social media platform that humans usually use for specific emotions for specific activities. This document collected Tweets multiple times and analyzed them using some machine learning algorithms such as Naïve Bayes, SVM, Random Forest Classifier, and LSTM and compared the results.

Sandeep Suri et al. [2020] Frustration is a type of emotion that stems from sadness or when an individual is dissatisfied with the product or service provided or the result of an experience or occasion that occurred in their presence. Buying products online is a new way to shop, using e-commerce or any virtual platform. Humans generally write reviews of a product that they have introduced and used. Surveys can be a great review, a bad review, or an impartial review. There may be a deep feeling of disappointment with the product purchased and used. Frustration is feeling disappointed, upset, cheated, or unable to get the other side's expected service. In this article, we attempted to uncover consumer frustration with the use of a machine learning algorithm in a dataset.

Kapil Sharma et al. [2019] the world is shrouded in many facts, and extracting the coveted records from such a huge amount of information is extremely complicated. The internet has the raw facts and turning those stats into insight, and they have to do the mining. Whenever a consumer performs any query on any search engine, the results come safely according to the importance of the file produced with the search engine. Certainly, the ranking is due to the algorithms they apply in the historical past. Most of the popular engines like Google like Yahoo, Bing, Google, and MSN, including their method of calculating ranking, and unique results are obtained in exclusive engines like Google for the same query due to the way the results are determined. Network pages differ from a set of rules to an algorithm. In this article, we tried to evaluate the internet page ranking algorithm as both strengths and weaknesses. This document demonstrates many of the most popular classification algorithms and tries to find a better-mixed answer to reduce the time spent searching on the page.

Asghar et al. [2018] among the many social media sites available, customers choose micro-blogging offers including Twitter for product offerings, social activities, and political trends. Twitter is an important source of information for sentiment rating software. Moderated and unsupervised devices that gain knowledge of technologies based primarily on analyzing Twitter statistics have been investigated in recent years, often resulting in the wrong kind of emotion. In this article, we recognize these issues and abandon a unified framework for categorizing Tweets using a mixed type system. The proposed approach aims to improve the performance of whole Twitter-based sentiment analysis structures by incorporating four classifiers: (a) the term classifier, (b) the code classifier, (c) the SentiWordNet classifier, and (d) the front-area-specific classifier stepped up the classifier. After taking advantage of the pre-processing steps, the input text is delivered through symbol and terminology classifiers. In the next row, SentiWordNet-based domain-specific workbooks are fully implemented for more precise text classification. Finally, emotions are categorized into a group of sentences and documents. The results found that the proposed method overcomes the limitations of previous methods by thinking of terminology, smileys, and terminology specific to a specific area.

Tripathy et al. [2017] Sentiment analysis is the essential branch of herbal language processing. It is the type of textual content intended to define the purpose of the text's author. The purpose can be liked (high quality) or critical (negative). This article assesses the consequences obtained using the Naive Bayes algorithm (NB) and the Support Vector Machine (SVM). These algorithms are used to classify an emotional overview that is either great or poorly rated. Classified the dataset considered for education and model testing in these panels based on the polar movie dataset, and an evaluation was made with the results that were present in the current literature for a vital screening.

Hajar et al. [2016] with the increasing use of smartphones, there is an increasing need for advanced capabilities that provide more brilliant smartphone user interaction. With the device provided, we intend to find user sentiments from their text exchange, manage the complexity of chat writing style and language development. We bear in mind that this type of device is the start of exciting packages that exploit users' emotional states. Our tool uses an unsupervised device learning algorithm that performs the emotion category based on a set of information generated from YouTube comments. This preference is intended for the similarity between YouTube comments and messages they instantly write fashionably. To classify a text entry into a specific emotion category, we calculated its similarity for each target emotion using the Pointwise Information Interchange Scale. Our approach yields an overall accuracy of ninety-two and 75%, which demonstrates the viability of our method.

Chopade et al. [2015] this article provides an overview of the growing field of sentiment detection from the textual content. It describes modern technology for detection techniques that typically fall into the following three main categories: Keyword-based recommendation strategies, which are entirely based on mastery and hybridization. The limitations of current sensing technologies are tested, and applicable solutions are recommended to enhance emotion-sensing skills in practical structures, emphasizing human-computer interactions. These responses consist of extracting key phrases with semantic evaluation and designing the ontology with the principle of evaluating feelings. Also, a case-dependent logical structure has been proposed to combine these solutions.

Munezero et al. [2014] one of the significant problems with computerized disclosure of influences on feelings, emotions, feelings, and reviews in the textual content is the lack of proper distinction between these subjective terms and knowledge of how they relate to each other. This impractical loss of differentiation leads to inconsistencies in the terms' use. Still, that was difficult to recognize the nuances expressed through the five terms, resulting in poor detection of phrases in the textual content. In light of this obstacle, this article explains the differences between these five subjective terms. That was revealed broad concepts of the computational language community for effective discovery and processing in the text.

Hoste et al. [2013] the achievement of suicide prevention, an important issue of concern for the public worldwide, is based on a sufficiently good assessment of suicide risk. Online systems are increasingly used to express suicidal thoughts; however, following directions is ineffective due to professionals' extra data. We examine whether recent advances in natural language processing, primarily in the exploration of sensations, can be used to correctly identify 15 distinct emotions, which are likely indications of suicidal behavior. The automatic emotion detection system has evolved into the use of dual auxiliary vector device classifiers. We hypothesized that lexical and semantic functions could be a suitable way of representing facts because feelings appear to be constantly lexical. The set of desirable characteristics for each of the unique emotions is determined by reconfiguring the primer. Spell checking was performed on the input data in an attempt to reduce the linguistic discrepancy. Overall performance rankings among feelings, with ratings of 68% and 86% of the F-rating. F-ratings of over 40% were made for six of the seven most common emotions: gratitude, guilt, love, actions, despair, and instructions. The most notable abilities are the word baggage of 3D shapes, logos and clues. Spell checking had a slightly nice effect on printing performance. We emphasize that high-quality computerized sentiment detection provides the advantages of classifier optimization and clarification of compound semantic lexical properties.

Kumar et al. [2012] Emotions can be expressed in many ways that they can be seen, including facial expressions, gestures, speech, and written text. Emotion detection in text documents is largely a content-based classification problem related to concepts from the fields of natural language processing and machine learning. This article mentions the reputation of emotions based on textual records and techniques used to discover emotions.

Agarwal et al. [2011] as network technology advances and grows, there are many facts out there on the Internet for Internet customers, and very few facts are also being generated. The Internet has become a platform for online domination, sharing of ideas and review sharing. Social networking sites such as Twitter, Facebook, and Google+ are gaining popularity as they allow people to rate and define their views on topics, chat with private communities, or post messages around the world. Much has been done in the area of sentiment analysis of the Twitter logs. This survey focuses specifically on assessing feelings for Twitter facts, which is useful for analyzing Tweets' records where opinions are particularly disorganized, heterogeneous, and of high quality, inadequate, or neutral in some cases. In this document, we present a survey and comparative analysis of current opinion polling strategies, such as the device to gain insight and primarily lexical operations, along with rating scales. Using various devices gaining knowledge of algorithms such as Naive Bayes, Max Entropy, and Support Vector Machine, we provide studies on Twitter data flows. We've also mentioned challenging situations extensively and Sentiment Analysis packages on Twitter.

Binali et al. [2010] Emotions are an essential component of a person's lifestyle and, among various things, have a

relative influence on decision making. Computers have been used in decision-making for some time. However, they have traditionally been reliable records of facts. Recently, researchers have discovered ways to uncover personal records used in blogs and other social networks online. This article presents theories of emotions that provide a basis for patterns of emotions. Suggest how these modes have been used when discussing computational techniques for detecting emotions. We advocate a fully hybrid architecture to discover emotions. The SVM algorithm is used to validate the proposed architecture and achieves a prediction accuracy of 96.43% in the network's weblog information.

With the analysis of the literature survey, we got some limitations in the previous works are listed below

3. Limitations in previous works

- Past work usually studied only one emotion classification. Controlling with various classifications concurrently not only allows performance similarities between various emotion categorizations on the same type of data.
- The primary focus of existed work was to support the study of Emotion Lexical Semantics, and consequently, no training data was presented

4. Proposed methodology

In previous studies of detecting emotions from text, many researchers have used specific strategies. In recent investigations, knowledge of the system has been used with every supervised and unsupervised workbook. Automatic knowledge is a better option for a larger data set, and classifier teaching is a more general approach than increasingly sentimental word dictionaries. To avoid the difficulties in the previous methodologies, in this paper Machine learning based supervised Naïve Bayes algorithm proposed to detect emotion relevant information from the twitter text. We implement the Naïve Bayes algorithm with 3, 5 and 10 cross validation times in our pre-processed text (Tweets and replies). We rate every text from NAVA and every clean text. We employed the classification on both the NAVA text and the whole clean text. The feature set used for Naïve Bayes incorporates words from each tweet and response. We employed the classifier on the same datasets twice – (i) to classify them according to their feelings and (ii) to classify them according to their corresponding feelings.

In this paper, we studied few other circumstances for our considerations. One person can change a different person in both positive and negative behaviors. Usually, when someone likes or re-tweets a tweet from another person, he/she consents to that tweet. But a person can reflect on another person's tweet and show either his/her support or conflict with that tweet. We examined this while measuring the impact based on all the circumstances – (i) the number of retweets, (ii) Number of comments that conform to the tweet or show similar sentiment/emotion, iii) Number of likes (iv) the number of comments that agree with the Tweet or show similar feelings / emotions, (v) the number of people observing the person

.

A) CLASSIFICTION TECHNIQUES

In the machine learning area, classification techniques have been established, which use different strategies to classify unlabeled data. Classification may want to request training information. Examples of known devices for classifiers include Naive Bayes, Maximum Entropy, and Support Vector Machine. They are classified as supervised system learning techniques, as they require training statistics. It is essential to note that effectively teaching the classifier will facilitate future predictions. In this approach we are using Naïve Bayes classifier to classify the emotion relevant text from the twitter, re-twitter text.

B) Naïve Bayes Classifier

Naive Bayes is a running technology that we use to build classifiers. Naive Bayes classifiers are simple probabilistic classifiers based on pi's theory while assuming a strong (naive) independence of most functions. In all naive Bayes classifiers, we assume that the price of a particular job is not biased towards the cost of some other characteristic, given the variable elegance.

The Naive Bayes classifier is based on Bayes' theory. Bay's theory is based on dependent possibility. Using Bayes' theorem, we can know the probability of event '**H'** occurring, given that B has happened. Event '**X'** is a guide, and event A is a view. The assumption here is that forecasters/opinions are unbiased. That is, the arrival of a certain property no longer affects the other.

Let X be a tuple of data. The P(H/X) is the subsequent prospect, or a posteriori prospect, of H trained on X. In distinction, $P(H)$ is the prior prospect, or a priori prospect, of H. Likewise, $P(X/H)$ is the posterior prospect of X trained on H. $P(X)$ is the first prospect of X. Bayes' theorem is valuable in that it offers a solution of evaluating the posterior prospect, $P(H/X), from\ P(H), P(X/H), and\ P(X)$.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

$$P(Y|x_1, \dots \dots, x_n) = P(y) \prod_{i=1}^{n} P(x_i|Y) \quad (2)$$

System architecture

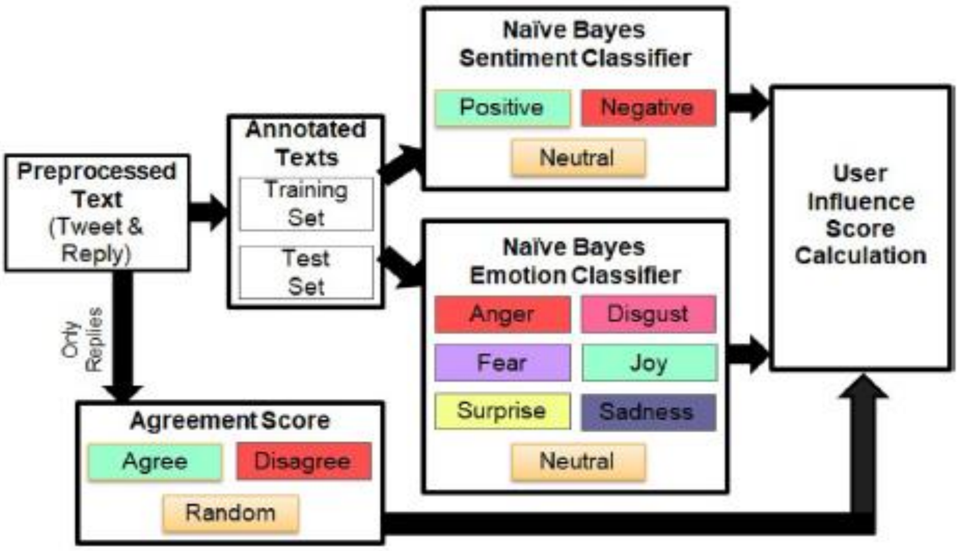

Fig.1 Proposed work architecture

As shown in the figure 1, shows a workflow of classification of preprocessed text and consideration of impact score to make a suggestion. They separated the preprocessed text into training and test sets for sentiment and emotion classifications. Then, satisfied emotion and sentiment into the estimate of "user influence score," which also managed the responses' compliance score. The impact score employed all these benefits and many tweets, retweets, followers, and followers to generate the final influence score. The recommender system took the score and developed recommendations as a list of users who shared similar emotions and sentiments on a specific topic.

5. Results and discusions

Several Machine Learning Algorithm has been implemented on data-set to recognize the emotion of user on review. On conclusion of our emotion detection pattern, for displaying the result, we used naive Bayes model

The proposed methodology conducted experiments on latest 500 tweets extracted from Twitter and proposed algorithm performance calculation performed using three evaluation metric of precision, recall, and Accuracy. These metrics are calculated using expressed using four parameters of True positive (TP), True negative (TN), False positive (FP), and False negative (FN). Each metric

True positive: It is an outcome where the model correctly predicts the positive class

True negative: It is an outcome where the model correctly predicts the negative class

False positive: It is an outcome where the model incorrectly predicts the positive class

False negative: It is an outcome where the model incorrectly predicts the negative class

The three metrics precision, recall, and F-score are followed by an equation given below

Every metric assisted with equation given below

$$Precision = \frac{True\ Positive}{True\ Positivee + True\ Negative} \times 100$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \times 100$$

$$Accuracy = \frac{No.of\ correctly\ labeled\ tweets}{Total\ No.of\ tweets\ in\ dataset} \times 100$$

Table.1 Emotion experimental classification

Emotion	Precision	Recall	Accuracy
Happy	88.1	89.2	93.1
Angry	87.2	89.5	93.1
Sad	85.6	88.2	93.1
Surprise	80.1	82.3	93.1
Fear	81.2	83.2	93.1

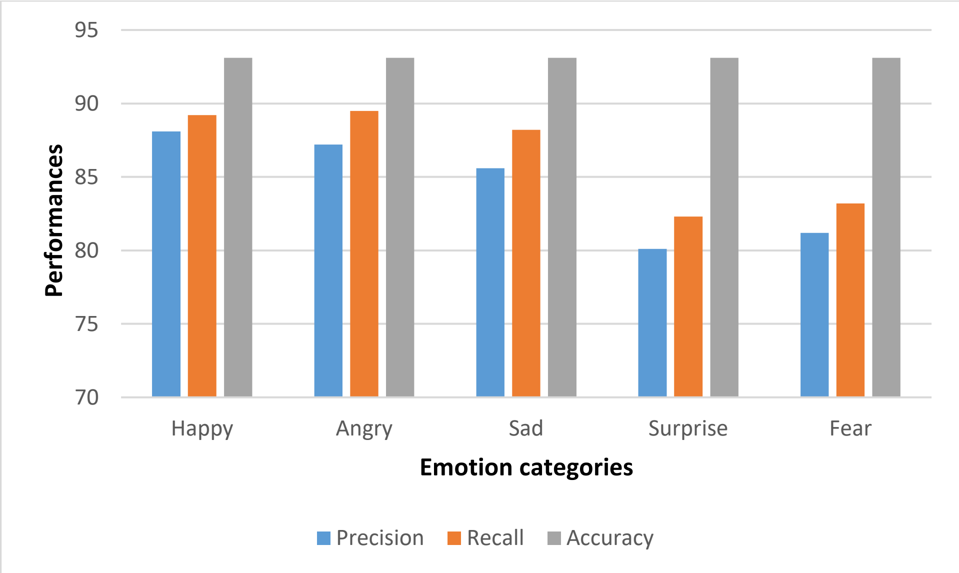

Fig.2 Emotion experimental classification

As shown figure 2, experimental taken between three emotion text such as sadness, joy, and surprise.

Table.2 Comparison between various classifications

Method	Precision	Recall	Accuracy
SVM	0.831	0.849	0.86
KNN	0.812	0.851	0.87
Decision Tree	0.950	0.961	0.89
Naïve Bayes	0.981	0.985	0.931

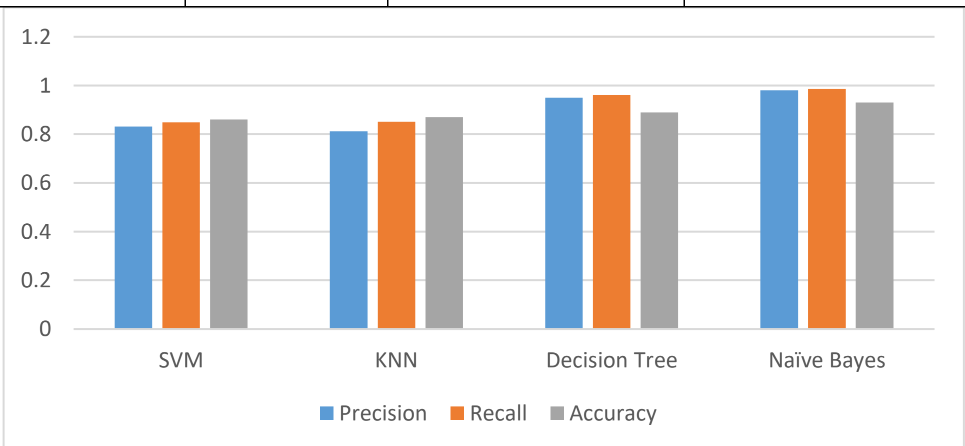

Fig.3 Comparison between various algorithms

Figure 3 shows the comparison between various algorithms in the classification of emotion relevant text from twitter text. Graph indicates that the proposed Naïve Bayes classifier got better precision, recall, and accuracy score when compared with previous classifiers of K-nearest neighbor (KNN) and Decision Tree (DT).

6. Conclusion

Emotion is an essential component of our daily lifestyle and behavior. The emotion shows how we can make decisions about the future, take another step to get going and start working in our daily lives. Most of the researchers applied various techniques to sentiment analysis and emotion recognition. Machine learning algorithms plays a vital role on the sentiment analysis. In this paper, the machine learning based Naïve Bayes (NB) classifier to predict the emotion relevant data from twitter data set. The proposed algorithm provides better performance in emotion text classification than traditional machine learning algorithms. The experimental results show that the proposed method provides better accuracy when compared with previous algorithms of KNN and DT.

References

1. Saif M. Mohammad: From once upon a time to happily ever after: Tracking emotions in mail and books. In: Decision Support Systems 53 (2012), H. 4, S. 730–741.
2. Seth A. Marvel, Steven H. Strogatz, 2011, “Continuous-time model of structural balance”,pp. 1771–1776.
3. Sheresh Zahoor, Rohilla Rajesh, 2020, “Twitter Sentiment Analysis Using Machine Learning Algorithms: A Case Study”,DOI. 10.1109/ICACCM50413.2020.9213011.
4. Sandeep Suri, Papneja Sachin, 2020, “Frustration Detection on Reviews Using Machine Learning”, IEEE, pp.1-5.
5. Kapil Sharma, Arushi Gupta, 2019, “Comparative Study of Ranking Algorithms.” pp. 73-77. IEEE.
6. M. Z. Asghar and F. Khan, 2018, "T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme,”.
7. Tripathy and S. K. Rath, 2017, “Classification of sentiment of reviews using supervised machine learning techniques,’ pp. 56-74, 2017.
8. M. Hajar, 2016, “Using YouTube comments for text-based emotion recognition’, Procedia Comput, pp.292–299.
9. C.R. Chopade, 2015, “Text based emotion recognition: a survey”, Int. J. Sci. Res. (IJSR) 4(6), pp.409–414.
10. Munezero M.D and J. Pajunen, 2014, “Are they different? Affect,feeling, emotion, sentiment, and opinion detection in text”, IEEE Trans. Affect.Comput, pp.101–111.
11. Hoste and B. Desmet, 2013, “Emotion detection in suicide notes”, pp.6351–6358.
12. S.N. Shivhare, S. Khethawat, Emotion detection from text, May 2012.
13. Agarwal and Rebecca J. 2011, "Sentiment analysis of twitter data.” pp. 30-38.
14. H. Binali, V. Potdar, 2010, “Computational approaches for emotion detection in text”, pp. 172–177.