EMOTION RECOGNITION BY TEXTUAL TWEETS CLASSIFICATION USING VOTING CLASSIFIER(LR-SGD)

[1]Dr. NASINA KRISHNA KUMAR, Ph.D., [2]BELLAMKONDA SAI HEMANTH

[1]Professor, Dept. of MCA, Audisankara College of Engineering and technology, Gudur.

[2]PG Scholar, Dept of MCA, Audisankara College of Engineering and technology, Gudur.

Received:
Accepted:

Keywords:
Text categorization, sentiment analysis, machine learning, opinion mining, emotion recognition, and artificial intelligence.

ABSTRACT

Opinion mining has become difficult due of the overabundance of user-generated information on social media. Twitter is used to gather opinions about products, trends, and politics as a microblogging site. Sentiment analysis is a method for examining someone's attitude, feelings, and views. Different individuals toward anything, and it is possible to do so by analysing tweets to determine how people feel about the news, regulations, social movements, and people. Opinion mining may be carried out without manually reading tweets by using Machine Learning models. Their findings could be useful to corporations and governments as they implement policies, programmes, and events. The use of seven machine learning models for emotion acknowledgment by dividing tweets into pleased and angry categories. An extensive performance comparison investigation revealed that the suggested voting classifier (LR-SGD with TF-IDF) gives the best results, with an F1 score of 81 and an accuracy of 79 percent. Two additional datasets, one binary and the other multi-class, were used to further confirm the suggested approach's stability and produce reliable findings.

1. INTRODUCTION

Recent advances in artificial intelligence have made automatic emotion identification, pattern recognition, and computer vision substantially more essential with applications in many different fields. In recent times, social media sites like Twitter have produced vast volumes of organised, unstructured, and semi-structured data. One of the most recent examples is the COVID-19 infodemic, which demonstrates how false information spread through social media can have a significantly greater impact and be far more catastrophic than a natural event like a pandemic. To correctly assign sentiment classifications on a broad scale, analysis is required. Accurate NLP methods and machine learning (ML) models for text classification are needed to complete these tasks. Twitter gives its users the chance to examine its data from a bigger and more comprehensive perspective. Due to the noisy nature of text input, efficient algorithms are crucial for automated labelling. Numerous research on the categorization of Twitter sentiment have been conducted in the past. Tweets are tiny posts that are sent over Twitter, which is a quick and effective microblogging platform that enables users to do so. Twitter is a popular social networking platform that is in high demand

worldwide. Twitter allows users to create free accounts, which have the potential to reach a huge audience. Twitter may be proven to be the ideal platform for business and marketing since it allows one to connect with extremely wealthy and well-known people, such as actors and celebrities, making their purchases appealing to both them and advertising. Every celebrity connects with fans andmay communicate with followers via Twitter. One of the best ways for lovers is
through such a site. It can type a post or a website link but only with a small note range (140 characters for each post). Free and available to the public, including adverts. The presence of personal ad clusters resembling those on other social networking websites is not problematic. It is rapid because the public who is following the relevant business will immediately get a tweet once it is posted on Twitter. Utilizing this source allows businesses and advertising to examine a variety of operational viewpoints that are quite important. They will get a quick reaction from their followers as a result of this. Surprisingly, a large number of firms enhance their transactions by increasing their Twitter following. Twitter helps its users by enabling them to discover new businesses, goods, and services, as well as websites, blogs, eBooks, and other things. As a result, Twitter users may click on links while pretending to do so and then naively invest in a manufactured good or review the products offered in order to participate in the profits. It is very simple to use since everyone can subscribe to get news and updates, businesses can tweet or retweet, they can choose their favourite or preferred recipients to send the tweets to, they can know how to promote the postings, and they can use it to endow their money and reputation. Super Bowls, Grammy Awards, Academy Awards, and other significant sporting and entertainment events use it to create a lot of talk across the world. On Twitter, competition between various products is increasing.

On social networks like twitter, people love to share their opinions about a certain product. Product owners are prepared to spend more money on social media platforms in order to promote their goods more effectively and increase sales. A person's experience with a product can assist the owner alter their marketing tactics, sales tactics, and product quality. Customer feedback is also given to owners or manufacturers through reviews. In order to categorise the consumer sentiment from the reviews given the volume of data created in this manner, a team of analysis experts is needed. It takes machine learning and ensemble learning classifiers to effectively classify the sentiment of the consumers because experts might make human errors in sentiment analysis. By classifying tweets using Tf and TF-IDF, this study evaluates several machine learning models for emotion identification. This study introduces a voting classifier (LR-SGD) and seeks to gauge how well-known ML classifiers perform on twitter datasets. The following are the main contributions:

- Support vector machine (SVM), decision tree classifier (DTC), naïve bayes (NB), random forest (RF), gradient boosting machine (GBM), and logistic regression are among the machine learning-based classifiers studied for the goal of recognising emotions (LR).
- Using LR and SGD together, a voting classifier (VC) created to categorise tweets surpassed TF-IDF.
- By using it on two separate datasets, one binary dataset (having hatred or non-hatred classes), and the other non-binary dataset, the suggested model stability is further confirmed.

The remaining parts of the essay are arranged as follows. Section II addresses the literature that is relevant to the ongoing study project.

The suggested technique is presented in Section III, along with a thorough explanation of the twitter dataset that was utilised for the experiment. Results are reported in Section IV, while Section V provides information on the stability of the suggested model. Section VI concludes the study project and offers ideas for additional research.

2. LITERATURE SURVEY

Corporations use sentiment analysis to determine the preferences of their customers for various goods, services, and brands. Additionally, it is crucial in evaluating data about firms and industries in order to reserve them for entity reviews. With the use of prototyping, Sarlan et al. [2] built a sentiment analysis, which categorised consumers' tweet-based opinions into good and negative ones. They separated their research into two parts. The first section is based on a literature review and uses current methodologies and methods for sentiment analysis. The second section describes the operations and requirements of the application before it is developed. In a different

study, Alsaeedi et al. [3] examined several sentiment analysis techniques used on the Twitter dataset and their results. The various methods and results of algorithm performance were contrasted. Approaches based on supervised machine learning, lexicon-based methods, and ensemble methods were applied. The authors employed four techniques, including Twitter sentiment analysis with supervised machine learning approaches and Twitter sentiment analysis with ensemble approaches. Lexicon-based approaches are used in Twitter sentiment analysis.

Many scholars have looked into vocabulary-based systems for classifying emotions. With the aid of domain-specific lexicon development, Bandhakavi et al. [4] carried out emotion-based feature extraction. They used a unigram mixture model to capture the relationship between word and emotion association. They attempted to categorise emotions using tweets with shoddy labels. Other cutting-edge strategies like Latent Dirichlet Allocation and Point wise Mutual Information fell short of the performance of their suggested design. Researchers use geo-related tweets to identify event-related tweets. They used particular tweets from one year's worth of regional celebrations. They also found other criteria that aided in the event finding process. Alsinet et al. [6] analyzed tweets from political domains. They claimed accepted tweets are stronger as compared to the rejected tweets. Rumor detection in tweets is performed by using an encoder to analyze human behavior in comments. To reduce the number of problems in the Twitter dataset, Hakh et al. [8] employed the SMOTE approach. In addition, they used a variety of feature selections to speed up the sentiment analysis process. The approach that the authors proposed, which was calculated alongside the dataset application choice, produced good outcomes for all operational assessment criteria.

They employed TF-IDF features to quantify the essential weight of phrases after applying pre-processing processes to their dataset. Next, classification techniques (such as AdaBoost, Linear SVM, Kernel SVM, Random Forest, Decision Tree, Naive Bayes, and K-NN) were utilised. Finally, accuracy and F1-score metrics were applied to relate the efficacy of classification. According to Sentiment's arrangement, Xia et al. established the proportionate training of the efficiency concerning collaborative approach in [9]. In the area of sentiment analysis, they established two categories of features. First of all, the feature set and word relations were entirely dependent on the portion of speech. Second, the well-known maximum entropy, support vector machine, and naive Bayes text categorization techniques.The third ensemble strategy was the fixed combination, followed by the weighted combination and metaclassifier combination. In addition to arena of Sentiment's configuration, they employed five widely used document-level datasets. The experiments conducted for this study demonstrate that ensemble approaches are more successful than other classifiers. Our search also demonstrates that an ensemble of two classifiers—logistic regression and stochastic gradient descent classifiers—provides superior results than other classifiers.

For the categorization of tweets and images, deep learning has been widely used by academics. A Tweets Classification for Sentiments from US Airline Companies was provided by Rustam et al. [12]. Pre-processing was done on the dataset by the researcher. The impact of feature extraction techniques, such as TF, TF-IDF, and word2vec, on classification accuracy has been investigated. Additionally, a specific dataset was used to study the operation of the long short-term memory (LSTM). A voting classifier (VC) is suggested in the researcher's paper to handle comparable elections. For determining outcomes, the voting classifier must rely on the spatial estimation (SE), stochastic gradient descent classifier (SGDC), and simple ensemble approach. Various ML classifier methods were assessed using working metrics such as precision, accuracy, recall, and F1-score. The findings show that the suggested VC is more effective than one of the phase actors. Students' performance increased as TF-IDF used a feature input, which was another demonstration of the effectiveness of machine learning.

A sentiment analysis of brief texts was looked at by Santos et al. [13]. In the experiment, researchers propose an original, deep convolution neural network that can analyse small texts' sentiments on a character-by-sentence basis. A sentiment study of customers of halal cuisine was reviewed by Mohamed et al. [14]. By looking at a strange case of 100,000 tweets relating to halal cuisine, our inquiry closes this gap. An expert designed lexicon of seed descriptors was used to guide the study. By examining how attitudes of halal cuisine are expressed on social networking websites, this investigation broadens and deepens

the conversation about a historically overlooked area. Investigation revealed generally good opinions of halal cuisine, and geo-located Twitter maps showed that "strict diaspora" largely uses digital presentations to spread the word about halal food. The sentiment analysis of a Twitter dataset using the NB algorithm was examined by Parveen et al. Analysts create a film informative collection using the Hadoop Framework that is accessible on the Twitter website as reviews, feedback, and opinions. Three classifications of sentiment—positive, negative, and neutral—are investigated using sentiment analysis on Twitter data.

SVM was examined by Alomari et al. [16] using TF-IDF. The study made available the Arabic-language Jordanian Twitter corpus, where Tweets are analysed for any positive or negative sentiment. It investigated unique directed machine learning opinion evaluation classifiers when applied to general themes discovered in either Modern Standard Arabic (MSA) or Jordanian language in the online lives of Arabic clients. A variety of weight plans, stemming, and N-grams terminology tactics, as well as scenarios, were examined through analysis. The Twitter benchmark dataset for Arabic Sentiment Analysis was created by Gamal et al. [17]. A benchmark Arabic dataset demonstrating social event approach concerning the most recent tweets in distinct Arabic dialects was proposed in an experiment for estimate study. More than 151,000 different assessments, divided into two categories (negative and positive), are included in the experiment dataset. SC employs ML algorithms, which are connected by learning arrangements. One basic methodology from an ML (lexicon-based approach) based approach is typically used to do sentiment analysis. The computations performed using SC on the dataset were 99.90% accurate using TF-IDF.

3. PRAPOSED SYSTEM

Numerous methodologies in ML have been applied in this study to achieve its goals. Various approaches and procedures were used to assess versatile experiments. The Voting classifier, which combines Logistic Regression with Stochastic Gradient Descent, performs better than all other ML models in terms of accuracy, recall, precision, and F1-score. The Twitter data collection that was utilised for this experiment was taken from the Kaggle repository. The dataset is initially pre-processed by removing irrelevant data. A training set and a testing set were created after dividing the data into two groups. A percentage of 70% was supplied to the training set whereas only 30% was given to the test set.

Following that, the training set is used to apply feature engineering approaches. The training set is used to train a number of machine learning classifiers, and the test set is used to evaluate them. Accuracy, recall, precision, and F1-score are the assessment criteria that were applied in this experiment.

3.1 Dataset

Many opposing tweets may be found in the dataset. The "Sentiment Analysis on Twitter data" dataset has 99989 entries in it. Each record is given an emotional polarity rating of either joyful or dissatisfied using the symbols 1 or 0. For the final dataset, English-language tweets are retained. Different characteristics may be found in the dataset. Features are included in Table 1 along with descriptions of each characteristic.

Table 1. Dataset Specifications.

Features	Description
Item ID	This is the index of record
Sentiment	This column contains sentiment happyand unhappy corresponding to tweets
Text	This column contains the textual tweets

3.2 Data visualization

Data Visualization helps to understand the hidden patterns lying inside the dataset. It helps to qualitatively get more details about the dataset by visualizing the characteristics of the attributes. Figure 1 shows the ratio of two target classes happy and unhappy. Figure 1 also illustrates that the happy class has more average than the unhappy class. Figure 1 displays the proportion of classes, and it reveals that 56.5 percent of tweets are happy tweets and 43.5 percent of tweets are connected to tweets that are sad.

3.2.1Data pre-processing

Datasets, which might be unstructured or semi-structured, include useless data in their raw form. Such pointless information.

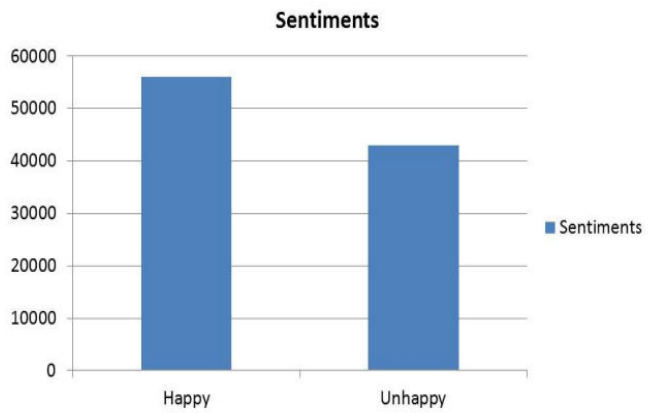

Figure 1. Countplot showing class-wise data distribution

Lengthens the model's training process and might harm its performance. Pre-processing is essential for boosting ML model effectiveness and conserving computational resources. The model's prediction accuracy is improved by text pre-processing. Pre-processing includes the following steps: tokenization, case conversion, stopword removal, and removal of numbers.

3.2.2 Feature extraction

The selection of features on a polished dataset comes after the crucial data pre-processing stage. Textual data in vector form is required for supervised machine learning classifiers. To take a training course on it. Textual elements are transformed. Tmploying TF and TF-IDF methods, into vector form. Techniques for extracting features assist in not only converting textual information into vector form but also in locating significant characteristics required for prediction. The majority of factors do not influence the prediction of the target class. The identification of tweets that are connected to being happy or upset is why feature extraction is crucial.

According to how frequently a word appears in the document, what does Term Frequency (TF) genuinely mean? TF is used to measure it. This is possible because each document varies in length, making a phrase appear much further in long papers than in small ones. Similarly to the mode on standardization:

$$TF(t) = \frac{\text{No.of times term t shows in a document}}{\text{Total no.of terms inside document}} \quad (1)$$

The phrase frequency is typically split with the length of the text (the total number of terms in the document). IDF: Inverse Document Frequency determines how important a phrase is inside the text. When TF is calculated, each phrase is given the same weight. However, it is acknowledged that words with strong conviction, such as "is,""of," and "that," might appear far more frequently than words with little conviction. Therefore, it is necessary to scale down common phrases while elevating unusual ones by computing the following:

$$IDF(t)=\log(e) \frac{\text{Total No.of documents}}{\text{No.of documents through term t in it}} \quad (2)$$

Data recovery uses term frequency (TF), which demonstrates how frequently a phrase or word appears in a report

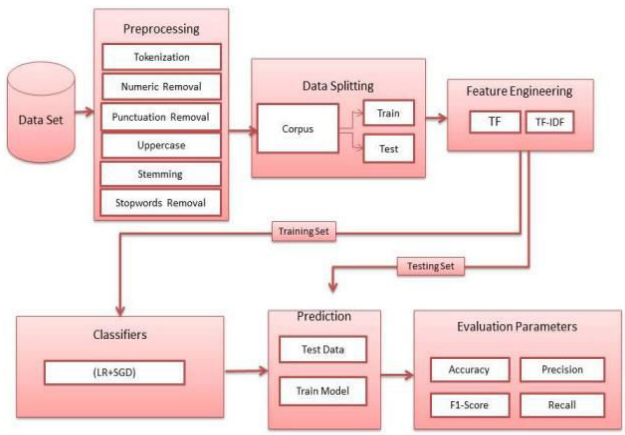

Figure 2: Proposed methodology architecture diagram.

3.3 Proposed models for tweets sentiment classification

Classifiers used to categorise tweets will be covered in this section. The anticipated data collection and work flow for this research project are shown in Figure 2. Support Vector Machines (SVM), Naive Bayes (NB), Random Forest (RF), Decision Tree (DT), Gradient Boosting model (GBM), Logistic Regression (LR), and Voting Classifier (Logistic Regression + Stochastic Gradient Descent classifier) were the five

supervised machine learning algorithms used in this study.

3.3.1 Random forest

RF is a tree-based classifier that uses a random tree-generation input vector. In order to build numerous decision trees and a forest, RF leverages random characteristics. After that, test data class labels are predicted using the combined vote of all trees. The decision trees with low value error are given higher weights. By taking into account trees with low error rates, overall prediction accuracy is increased.

3.3.2 Support vector machine

Support vector machine (SVM) execution as sentiment analysis is understood. SVM characteristics preference, restricts and employs the methods for evaluation, and evaluatesrecords obtained inside the index area. Every magnitude's arrangements of vectors include essential information. To accomplish this goal, information has been grouped in type (presented as a vector). The boundary is then stratagem-classified in two training sessions.

3.3.3Naive Bayes

The Bayes Theorem underlies the Naive Bayes(NB) ordering strategy, which relies on strong (naive) independent assumptions across stabilities. The NB classifier assumes that a particular class element's proximity is limited to the proximity of a few different factors. For instance, a natural organic product is likely to be recognized as an apple if it is dark red in colour, round, and about three creeps in diameter. Naive Bayes classifiers are a collection of fundamental "probabilistic classifiers" used in machine learning that apply Bayes' hypotheses with gullible opportunity assumptions between the features. These Bayesian network models are thought to be the least troublesome ones.

3.4 Decision tree

The supervised ML technique known as the DT algorithm is frequently employed in problems involving regression and classification. The key issue, known as attribute selection, is choosing the root node of a tree at each level. The approaches for attribute selection most frequently employed are information gain and the Gini index. In this study, the Gini index is utilised to determine the likelihood of a root node by squaring the attribute value total and then subtracting 1 from the result.

3.4.1Gradient boosting machine

For regression and classification problems, the GBM is a popular ML-based boosting model that uses an ensemble of weak prediction models, often decision trees. During boosting, weak students are transformed into strong students. Every tree that is formed is a modified version of the one before it and uses gradient as the loss function. Loss determines how well a model's coefficients fit the underlying data. Model optimization employs a logical loss function.

3.4.2 Logistic regression

In LR, class probabilities are calculated based on output, therefore they may be used to predict whether an input belongs to class X or class Y with a probability of x or y. Predicted output class is X if x exceeds y; else, it is Y. Insight is a logistic strategy that shows the likelihood of a specific group occurring or not, such as top/bottom, white/black, up/down, positive/negative, or happy/unhappy.To determine if a picture contains a snake, dog, deer, etc., each item famous in the image would be assigned a probability somewhere in the range of 0 and 1 with whole addition to one. This may be extended and used to show a limited number of classes concerning occurrences.

3.4.3 Stochastic gradient descent

Stochastic Gradient Descent is one of the varieties of gradient descent (SGD). A goal work with the required perfection qualities, such as differentiable or sub differentiable, can be advanced using the iterative SDGD technique. It calculates the degree of progress in light of the evolution of alternative variables. Since it substitutes the actual angle, which is found from the entire informative index, with a gauge of it, which is obtained from a randomly

selected subset of the information, it is quite acceptable to interpret it as a stochastic guess of inclination plummet progression.

3.4.4Voting classifier

Voting Classifier (VC) is a cooperative learning method that combines the predictions of numerous individual classifiers in order to achieve higher

performance than a single classifier [34]. It has been demonstrated that a combination of several classifiers may be more effective than any single one. The VC is a meta-classifier used to combine equal or fictitiously exceptional ML classifiers for order through greater part tossing a vote form. It conducts both "hard" and "soft" voting. Hard voting provides the researcher the opportunity to predict the class name instead of the most recent class mark, which has frequently been predicted by characterization models. Through averaging the class-probabilities, soft voting gives researchers the opportunity to predict the class names.

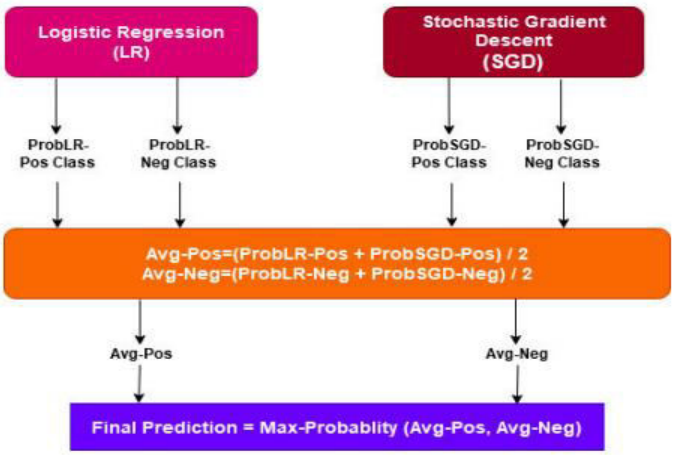

Figure 3:Proposed voting classifier architecture (LR-SGD)

By using the sigmoid function on the input for binary classification, LR determines the posterior probability p(Ct|v). VC is best described as:

$$p^{\wedge} = \text{argmax}\{\sum_i^n \boldsymbol{LRi}, \sum_i^n \boldsymbol{SGDi}\} \quad (3)$$

After that, the probabilities for each test example by both LR and SGD passes through the soft voting criteria as shown in Figure 3.

An explanation of the VC's functions may be found in example. Whenever a particular sample goes through the LR With SGD, each class is given a likelihood score (that either a good or bad thing). Let the probability rating of LR be For ProbLR Pos and ProbLR Neg, respectively, 0.966, 0.024, and the probability score for classes and SGD is 0.997, 0.002 for ProbSGD is Pos, whereas ProbSGD is Negative. Next, the It is possible to compute the average probability for the two classes as

$$\text{Avg -Pos} = (0.966 + 0.997)/2 = 0.9815$$

$$\text{Avg-Neg} = (0.024 + 0.002)/2 = 0.013$$

The maximum probability is given by MaxProb(AvgPosandAvgNeg). The affirmative class is present in this sample response. Since the real class in the dataset is "positive" and the projected class is, too. To reach a judgement, the suggested VC aggregates the projected probabilities from both classifiers. MLR and MSGD, which based on the dataset, and then calculate the likelihood of both different classifications. An average likelihood is determined for two classifiers forecast the probability for each class. The final class of the decision function is then determined. Review that is based on the class's highest average probability. The LR-mechanism SGD's of operation is seen in Algorithm 1.

3.5 Evaluation metrics

ML models are assessed using a variety of widely used performance metrics, including accuracy, recall, and precision.

Algorithm 1: Stochastic Gradient Descent and Logistic Regression Assembled (LR-SGD)

Input: input data $(x, y)_{i=1}^{N}$
MLR = Trained_ LR
MSGD = Trained_ SGD
1:**for** *i=1to M* **do**
2:**if** *MLR* ≠0 &*MSGD* ≠0 &*training_set* 6= 0 **then**
3: *ProbSGD − Pos = MSGD.probibility(Pos − class)*
4: *ProbSGD − Neg = MSGD.probibility(Neg − class)*
5: *ProbLR − Pos = MLR.probibility(Pos − class)*
6: *ProbLR − Neg = MLR.probibility(Neg − class)*
7: Decision function $=max(\frac{1}{N_{classifier}}\sum classification$ (*Avg(ProbSGD−Pos,ProbLR−Pos, Avg(ProbSGD−Neg,ProbLR−Neg))*)
8: **end if**
9: Return fifinal labelb *p*^
10:**end for**

F1-score in activities requiring categorization. Accuracy is a metric for measuring how accurately a forecast was made.

$$\text{Accuracy} = \frac{\text{No.of times term t shows in a document}}{\text{Total no.of terms inside document}} \quad (4)$$

While incase ofbinary classification, accuracy is measured as:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

While TP stands for "true positive," FP for "false positive," TN for "true negative," and FN for "false negative" [10].

TP : TP denotes an accurately predicted class's optimistic predictions.

FP : FP indicates the unfavourable forecasts of an inaccurate anticipated class.

TN : TN stands for an accurately anticipated class's unfavourable forecasts.

FN :FN represents the positive predictions of a incorrectly predicted class.

The percentage of positively classified tuples that are genuinely positive is determined by precision, which determines how accurate a classifier is. It may be calculated as:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

Recall, on the other hand, gauges comprehensiveness and displays the proportion of accurately labelled true positive tuples. Recall is measurable as:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

Accuracy by itself is insufficient as a criterion for evaluation for the imbalance dataset. In these circumstances, the F1 score, which is the harmonic mean of recall and accuracy, can be useful. It carries out statistical analysis and determines the score, which ranges from 1 to 0.

Models	Accuracy	Precision	Recall	F1-Score
RF	74%	74%	79%	77%
SVM	76%	76%	80%	78%
NB	75%	75%	78%	75%
DT	74%	74%	77%	76%
GBM	74%	72%	79%	76%
LR	76%	79%	82%	80%
VC(LR-SGD)	78%	78%	84%	81%

Table 2. Results of classification using TF features for all machine learning models.

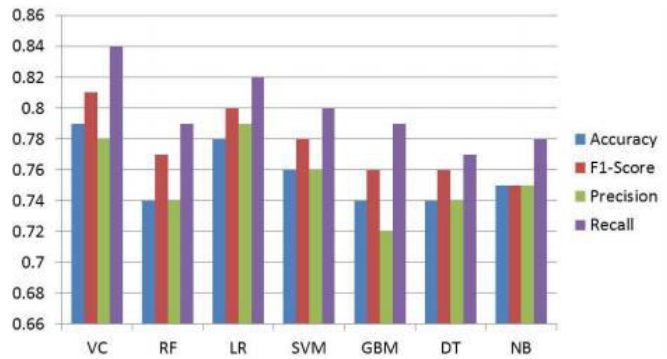

Figure 4: A comparison of all machine learning models' classification outcomes utilising TF characteristics.

Models	Accuracy	Precision	Recall	F1-Score
RF	74%	74%	79%	77%
SVM	76%	76%	80%	78%
NB	75%	75%	75%	78%
DT	74%	74%	77%	76%
GBM	74%	72%	79%	76%
LR	78%	79%	82%	80%
VC(LR-SGD)	79%	78%	84%	81%

Table 3. Classification result of all machine learning models using TF-IDF features.

Taking into account the model's recall and accuracy. The F1-score is calculated as:

$$\text{F1score} = 2\frac{\text{precision.recall}}{\text{precision+recall}} \quad (8)$$

RF = Random Forest Algorithm
SVM = Support Vector Machine
NB = Navie Bayes
DT = Decision Tree
GBM = Gradient boosting machine
LR = Logistic regression
SGD = Stochastic gradient descent
VC = Voting classifier

4. RESULTS AND DISCUSSION

The specifics of the experiment carried out for this study are provided in this part, along with a discussion of the findings. TF and TF-IDF features are used to evaluate classification methods. The most accurate voting classifier is one that combines stochastic gradient descent and logistic regression. The accuracy, recall, precision, and F1-score of classification using TF characteristics are shown in Table 2.

Table 3 displays the TF-IDF technique's classification accuracy, recall, precision, and F1 score. The voting classifier's accuracy score of 79 percent was higher than LR's score of 78 percent. LR attained the maximum level of accuracy.

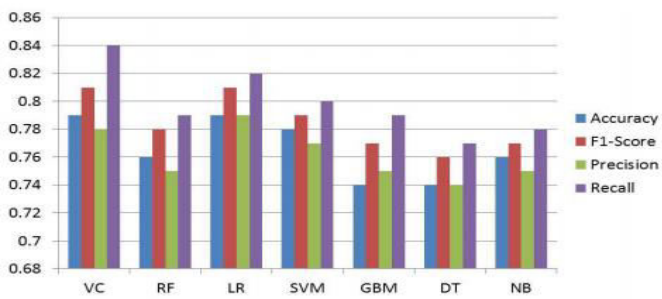

Figure 5: A comparison of all machine learning models classification outcomes utilising TF-IDF characteristics.

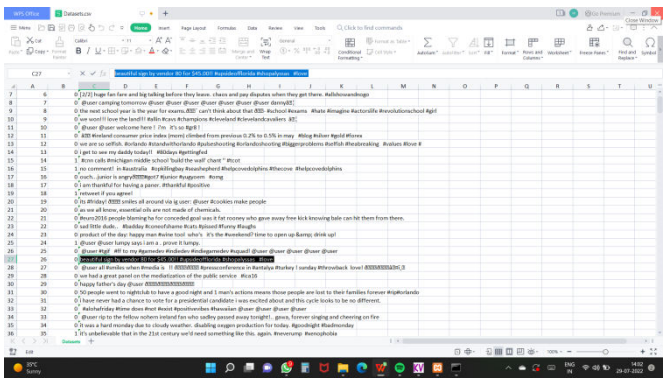

Figure 6. Dataset

In the figure, we are giving the data from the dataset.

Figure 7. Home page

In the above figure, we are able to login to the home page.

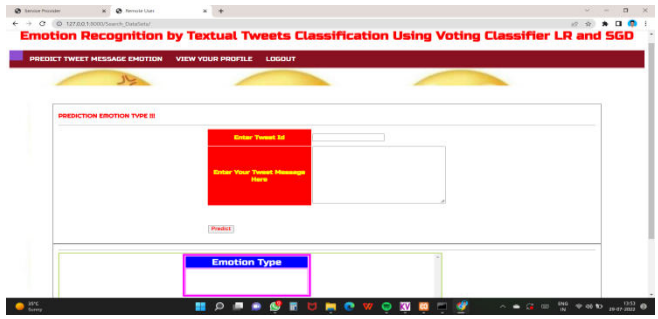

Figure 8. prediction page

In the above figure, we are entering the tweet id and tweet text. Finally predict the emotion type.

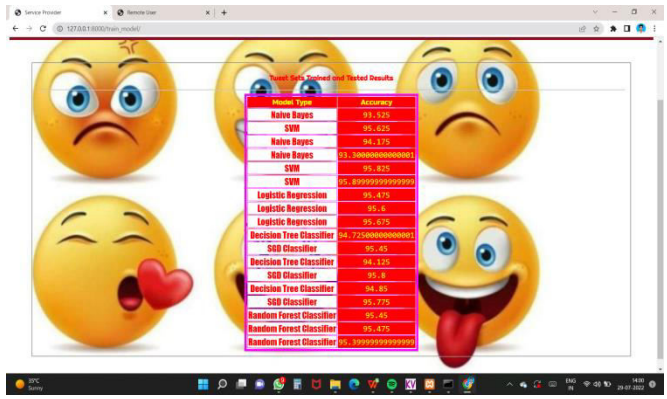

Figure 9. Accuracy levels in all the ML models

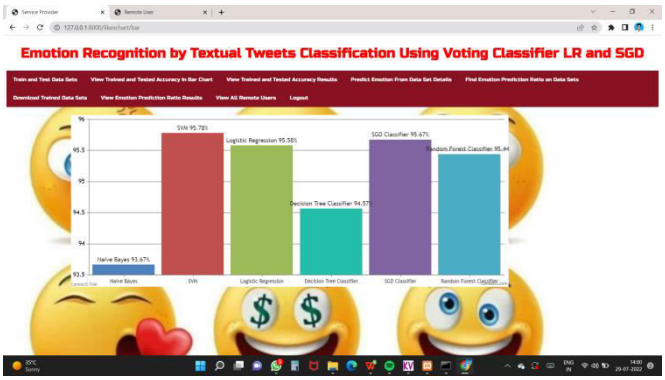

Figure 10. Accuracy levels in bar chart

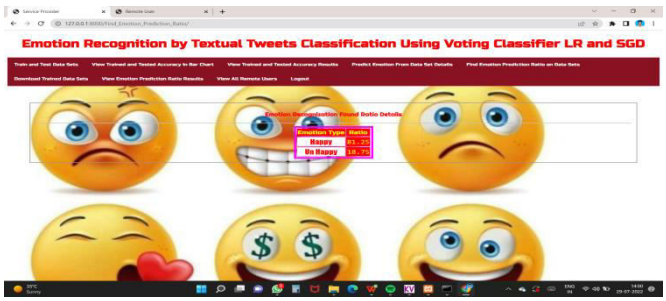

Figure 11. Final ratio

In the above figure, depicts the ratio about the happy and unhappy.

5. STABILITY OF THE PROPOSED MODEL

From the findings presented above. All other conventional models are outperformed by the proposed All typical machine learning-based models did not perform well on all three datasets, as can be seen Voting Classifier ensemble. If the cause of RF's poor performance is investigated, particularly with regard to dataset 2, it is found that RF is an ensemble approach that consists of merging many trees to cope with outliers and noise. However, it might be challenging to understand the relationships in the input data for a big dataset [42].

SVM works by dividing classes with the use of a hyperplane and perform well when dealing with binary classification issues. By creating hyperplanes across classes, it separates class labels, although SVM is typically unable to separate the data for multiclass situations. On all datasets, SVM outperformed the majority of the conventional ML models, including RF, GBM, and NB. This work used a mix of ML models as voting classifiers to address the shortcomings of ML models. The suggested VC(LR-SGD) outperforms traditional ML-based models on all datasets, as can be observed in tables 3.

6. CONCLUSION AND FUTURE WORK

By categorising tweets as happy or unhappy, this article presented a unique LR and SGD combination to be used as a voting classifier for emotion identification. Our tests demonstrated that efficient pattern recognition and effective model combination averaging may be used to raise the performance of models. Seven machine learning models are put to the test in experiments: (1) SVM, (2) RF, (3) GBM, (4) LR, (5) DT, (6) NB, and (7) VC (LR-SGD). Two feature representation approaches, Tf and TF-IDF, were also used in this work. All models on the tweet dataset performed well, according to the findings, however our suggested voting classifier, VC(LR-SGD), outperformed them all by using both TF and TF-IDF.The proposed model, which has 79 percent Accuracy, 84 percent Recall, and 81 percent F1-score, produces the best results when utilising TF-IDF. The suggested model was subsequently tested on two more datasets, and it produced reliable outcomes. In order to increase performance, future study will evaluate more feature engineering methodologies and investigate more ensemble model combinations. Additionally, fresh approaches of dealing with snarky remarks will be researched.

REFERENCES

[1] N. F. F. da Silva, E. R. Hruschka, and E. R. Hruschka, 'Tweet sentiment analysis with classifiifier ensembles,'' *Decis. Support Syst.*, vol. 66, pp. 170–179, Oct. 2014.

[2] C. Kariya and P. Khodke, ''Twitter sentiment analysis,'' in *Proc. Int. Conf. Emerg. Technol. (INCET)*, Jun. 2020, pp. 212–216.

[3] A. Alsaeedi and M. Zubair, ''A study on sentiment analysis techniques of Twitter data,'' *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 2, pp. 361–374, 2019.

[4] A. Bandhakavi, N. Wiratunga, D. Padmanabhan, and S. Massie, ''Lexicon based feature extraction for emotion text

classifification,'' *Pattern Recognit. Lett.*, vol. 93, pp. 133–142, Jul. 2017.

[5] J. Capdevila, J. Cerquides, J. Nin, and J. Torres, ''Tweet-SCAN: An event discovery technique for geo-located tweets,'' *Pattern Recognit. Lett.*, vol. 93, pp. 58–68, Jul. 2017.

[6] T. Alsinet, J. Argelich, R. Béjar, C. Fernández, C. Mateu, and J. Planes, ''An argumentative approach for discovering relevant opinions in Twitter with probabilistic valued relationships,'' *Pattern Recognit. Lett.*, vol. 105, pp. 191–199, Apr. 2018.

[7] W. Chen, Y. Zhang, C. K. Yeo, C. T. Lau, and B. S. Lee, ''Unsupervised rumor detection based on users' behaviors using neural networks,'' *Pattern Recognit. Lett.*, vol. 105, pp. 226–233, Apr. 2018.

[8] H. Hakh, I. Aljarah, and B. Al-Shboul, ''Online social media-based sentiment analysis for us airline companies,'' in *New Trends in Information Technology*. Amman, Jordan: Univ. of Jordan, Apr. 2017.

[9] R. Xia, C. Zong, and S. Li, ''Ensemble of feature sets and classifification algorithms for sentiment classifification,'' *Inf. Sci.*, vol. 181, no. 6, pp. 1138–1152, Mar. 2011.

[10] M. Umer, S. Sadiq, M. Ahmad, S. Ullah, G. S. Choi, and A. Mehmood, ''A novel stacked CNN for malarial parasite detection in thin blood smear images,'' *IEEE Access*, vol. 8, pp. 93782–93792, 2020.

[11] S. Sadiq, A. Mehmood, S. Ullah, M. Ahmad, G. S. Choi, and B.-W. On, ''Aggression detection through deep neural model on Twitter,'' *Future Gener. Comput. Syst.*, vol. 114, pp. 120–129, Jan. 2021.

[12] F. Rustam, I. Ashraf, A. Mehmood, S. Ullah, and G. Choi, ''Tweets classifification on the base of sentiments for US airline companies,'' *Entropy*, vol. 21, no. 11, p. 1078, Nov. 2019.

[13] C. D. Santos and M. G. D. Bayser, ''Deep convolutional neural networks for sentiment analysis of short texts,'' in *Proc. 25th Int. Conf. Comput. Linguistics*, Aug. 2014, pp. 69–78.

[14] M. Mohamed, ''Mining and mapping halal food consumers: A geo-located Twitter opinion polarity analysis,'' *J. Food Products Marketing*, vol. 24, pp. 1–22, Dec. 2017.

[15] H. Parveen and S. Pandey, ''Sentiment analysis on Twitter data-set using naive Bayes algorithm,'' in *Proc. Int. Conf. Appl. Theor. Comput. Commun. Technol.*, Jan. 2016, pp. 416–419.

[16] K. M. Alomari, H. M. Elsherif, and K. Shaalan, ''Arabic tweets sentimental analysis using machine learning,'' in *Proc. Int. Conf. Ind., Eng. Appl. Appl. Intell. Syst.*, Jun. 2017, pp. 602–610.

[17] D. Gamal, M. Alfonse, E.-S. M. El-Horbaty, and A.-B. M. Salem, ''Twitter benchmark dataset for arabic sentiment analysis,'' *Int. J. Modern Edu. Comput. Sci.*, vol. 11, no. 1, pp. 33–38, Jan. 2019