

# A Machine Learning-Based Technique with Intelligent WordNet Lemmatize for Twitter Sentiment Analysis

S. Saranya\* and G. Usha

Computing Technology, SRMIST, Chennai, Tamil Nadu, India

\*Corresponding Author: S. Saranya. Email: saranya.it78@gmail.com

Received: 02 May 2022; Accepted: 07 June 2022

**Abstract:** Laterally with the birth of the Internet, the fast growth of mobile strategies has democratized content production owing to the widespread usage of social media, resulting in a detonation of short informal writings. Twitter is micro-blogging short text and social networking services, with posted millions of quick messages. Twitter analysis addresses the topic of interpreting users' tweets in terms of ideas, interests, and views in a range of settings and fields. This type of study can be useful for a variation of academics and applications that need knowing people's perspectives on a given topic or event. Although sentiment examination of these texts is useful for a variety of reasons, it is typically seen as a difficult undertaking due to the fact that these messages are frequently short, informal, loud, and rich in linguistic ambiguities such as polysemy. Furthermore, most contemporary sentiment analysis algorithms are based on clean data. In this paper, we offers a machine-learning-based sentiment analysis method that extracts features from Term Frequency and Inverse Document Frequency (TF-IDF) and needs to apply deep intelligent wordnet lemmatize to improve the excellence of tweets by removing noise. We also utilise the Random Forest network to detect the emotion of a tweet. To authenticate the proposed approach performance, we conduct extensive tests on publically accessible datasets, and the findings reveal that the suggested technique significantly outperforms sentiment classification in multi-class emotion text data.

**Keywords:** Random Forest; sentiment analysis; social media; term frequency and inverse document frequency; twitter; wordnet lemmatize

## 1 Introduction

These days, people are more likely to express their views on social media establishes such as Instagram than they were a few years ago; they are also more likely to use instant messaging programmes like WhatsApp and Skype than they used to be. Here you'll find suggestions for new items as well as real-time feedback from users [1]. The rise of these messages has given researchers novel avenues for investigating public opinion via social media. Real-time material has also been used by businesses to learn more about their customers and devise relevant tactics. Use this data to either remedy a product's shortcomings or observe people's sentiments before introducing a fresh product [2]. Social media posts



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

can be categorised as either positive or negative based on their content. This is why “your curly and glossy hair appear fantastic” is likely to elicit good feelings. This statement, thus, has an uplifting effect. However, attempting to decipher the feelings or emotions expressed in a social media post using technology is a difficult endeavour.

In the microblogging and social networking site Twitter, huge instant messages (i.e., tweets) are posted every single day. In order to process and search tweets, researchers have taken use of the limited amount of characters (maximum of 140 characters for each tweet) and the use of hashtags between phrases. [3] Sentiment analysis of Twitter feeds is now a prominent research and business activity. An analysis of the opinions expressed in the tweets is addressed by Twitter Sentiment Analysis (TSA) [4]. A suitcase problem, sentiment analysis covers a variety of Natural Language Processing (NLP) tasks, with sarcasm discovery and the extraction of aspects. Our attention is on the polarity identification task, which is a subtask of sentiment analysis.

In the literature of TSA, a lot of effort has been put into achieving an effective representation of tweets using supervised machine learning algorithms. A variety of features have by this time been projected, ranging from simple n-gram representations through meta-level features to word embeddings [5]. It may be impractical to search for and identify tweets that express favourable or negative sentiment manually, given the large volume of tweets released each day. In this context, sentiment analysis is known as opinion mining. One of the most pressing issues in this subject is the spontaneous detection of ideas and emotions conveyed in brief informal writings like tweets. Sentiment analysis might be difficult because of tweets’ casual language style, misspelling words, and sloppy use of grammar.

The use of slang, non-standard punctuation, and unique phrases in social media postings necessitates a sophisticated pre-processing framework in order to comprehend the messages’ abbreviations and misspelt words. Tweets should be harmonised in terms of polysemy, syntax, semantics, and word sentiment information in order to progress sentiment analysis. In semantic analysis, tweets’ semantic information and distributed word representation have been frequently used by Word2Vec and GloVe [6,7]. However, the polysemy and noise in tweets cannot be taken into account by these methods. Words like “good” and “bad” have different meanings depending on the context in which they’re used. It is impossible to deal with polysemy with traditional word embeddings, which assign the same representation to words regardless of context and meaning. As a result, the meaning of words such as “good” or “bad” is lost when they are represented using typical word embedding methods. Due to the unstructured and low quality of content, and out of vocabulary (OOV) terms [8] all severely impact sentiment analysis.

Polysemy in context, complex properties of words (such as syntax, semantics and sentiment), and concerns linked to OOV words are some of the obstacles that this work addresses. Finally, we use the **Radio Frequency (RF)** to classify sentiment and TF-IDF to extract features from the data set [9]. We demonstrate that the proposed approach outperforms the current state of the art by employing a collection of tweets including emotional content. The following are the study’s most important contributions:

A smart twitter pre-processor that reduces noise by conducting spell correction, word segmentation, and normalisation is proposed [10].

As a first step, we describe an algorithm that uses machine learning and Latent Dirichlet Allocation to address language uncertainty, discover the deep relationships between words [11].

In demonstrate the effectiveness of the proposed outline, we conduct extensive trials on a dataset of emotion-related tweets.

The rest of the paper is laid out in this way: The existing works are summarised in Section 2. The structure of the proposed framework is labelled in Section 3. The framework is evaluated and analysed in Section 4. Conclusions are included in Section 5.

## 2 Literature Review

In this setting, recent research has attentive on the use of sentiment examination in social media marketing and financial forecasts [12]. As an open research question in the financial market, Li et al. [13] recently researched how to association technical needles from stock prices with news opinions from textual news items. A two-layer **Long short-term Memory (LSTM)** network was built for this purpose, using sentiment dictionaries to represent news sentiment, and an LSTM network to forecast stock prices. Both technical indicators and current news attitudes were found to beat the baseline models that only used one at a time.

To address the issue of balancing the trade-off amongst asset returns and investor risk, Xing et al. [14] looked at the difficulty of including public mood into the problem of asset allocation. LSTM neural networks and an evolving clustering approach (ECM) were used to codify sentiment information in financial markets. Thus, they offered to compute sentiment from media by the Septic framework, saying that it is capable of sentiment examination not only at the document or phrase level, but also at the conceptual level [15].

When it comes to sentiment analysis, Xing et al. [16] looked into how a basic translation technique could be used to solve the problem. When it comes to analysing the sentiment of a text, machine translation schemes like Google Translate or Microsoft Translator Text **Applicatrion Programming Interface (API)** can create accurate translations that can be utilised for this purpose. Unsupervised reference solutions are used in semi-supervised learning, as described by Araújo et al. [17]. Using random projections and SVM, they developed a unique semi-supervised ideal for the task of emotion appreciation.

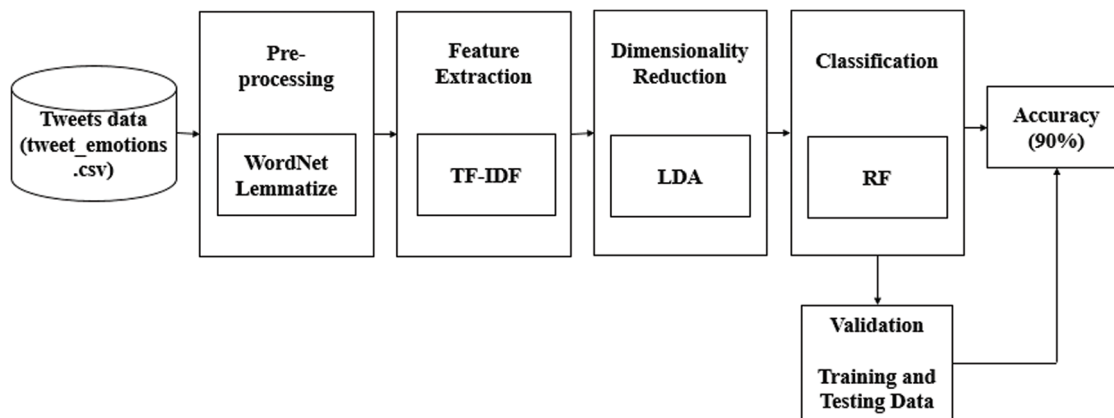
Semi-supervised approaches to sentiment analysis are already appearing in the literature, taking advantage of the vast amounts of unlabelled data now available as a result of the explosion of social media networks. In order to perform the aspect-level sentiment, the author from [18] developed the Semi-supervised Aspect Level Sentiment Classification Model based on Variational Autoencoder model. The encoder and decoder is filled with the input of given aspect for classification and done the aspect level classification. But, the lexical ambiguity is ignored in this framework for TSA.

Classifier ensembles have recently been presented by Emadi et al. [19] in the field of Twitter sentiment sorting. A combination of unigrams and bigrams is supplied to each of the three ML algorithms, such as Support vector Machine (SVM), NB, and ME as the basic classifier. To supplement these classifications, an NLP-based technique is employed. In order to pick approaches that complement one another, the classifiers are chosen based on diversity measures. A learning fusion technique is used to give a polarity orientation to each tweet once a varied set of classifiers, i.e., classifiers with adequate diversity, has been identified. Each classifier's choice is combined using the Choquet Fuzzy Integral (CFI) approach as a meta-learning strategy. This sentiment analysis framework, on the other hand, utilised a reverse polarity method to deal with negation. For tweet sentiment analysis, Jianqiang et al. [20] discovered that the findings can be improved by incorporating DCNNs into the analysis of short texts and reporting that this technique can improve sentiment analysis accuracy. However, the solutions listed above do not take into account situations involving polysemy. To create vector representations of words for the sake of sentiment analysis, word embedding is a common technique to use. For aspect-level sentiment analysis, Ray et al. [21] integrated a deep learning method with rule-based algorithms. Reverse-polarity negation handling was employed by them.

## 3 Proposed Methodology

The foremost objective of this opinion mining process is to analyse the positive, negative and neutral opinions about a product or a service posted by people, group or society. This research work uses Random Forest (RF) [22] for analyzes these tweets. RF classification effectively reduces the size of the duplication problem that arises when creating convenient defect classification boundaries. In adding, the

classification attitude speeds up the training process while preserving the accuracy of the competition classification. Fig. 1. displays the formation of Twitter's emotion classification using proposed model.



**Figure 1:** Working flow of proposed methodology

Classifying emotions may be done quite effectively using this paradigm. In this framework, emotions are divided into four steps of data gathering, Twitter pre-processing, which includes filtering to remove Twitter's unique traits. TF-IDF [23] functions are then used to classify tweets based on their behaviours. The LDA [24] is initialised to categorise tweets, and these categorizations are used as input for RF to classify tweets.

### 3.1 Data Collection

To convey one's thoughts and feelings, text remains the most popular method of communication. Textual data makes extracting emotions from it even more challenging. Emotions can be detected more easily in literature if the words that express the specific feeling are clearly defined. Emotion is often expressed in a more subtle manner. In some cases, a single piece of writing may contain many emotions. Then, there are emotions and words that are unclear, some words have numerous meanings, and multiple words mean the same emotion in different contexts in some texts. Slang and sarcasm are used in some of the material [25]. Multilingual language, spelling problems, acronyms, and grammatically incorrect sentences are some of the more typical aspects of internet communications. Textual data's constraints can make it difficult, if not impossible, to detect emotions automatically. That's why we're using a dataset from <https://kaggle.com/pashupatigupta/emotion> detection from text as input data for multiclass text classification. Collection of tweets with emotions annotated is what we have here. Tweet ID, emotion, and content make up the three columns. The tweet is included in "content." The tweet's "sentiment" conveys the tweeter's feelings. Concern accounts for 22% of the feeling, while neutrality and other emotions account for 21% and 57%, respectively. There are 39827 unique values in this multi-class text data. Text data with multi-class classifications are described in the experimental part [26].

### 3.2 Pre-Processing

The initial phase in data mining is pre-processing, which is the simplest and most fundamental. This is a crucial stage in the process of extracting data from a database. It's for this reason that a pre-processing phase is performed before the extraction of features [27]. The data is pre-processed after it has been collected from the dataset. Because the Twitter language includes several unique properties that may not be essential to the categorization process [28], such as usernames [29], links [30], and hashtags, tweets must be pre-processed before classification. When cleaning data, lexical analysis is used. Different word forms are broken down into

their most basic form. In the input clause, post is labelled [28]. On the basis of stochastic and rule-based approaches, the speech tag algorithm is built (e.g., the E. Brill tagger). Then grouping the same words into a single word is called lemmatisation. An informal word is morphed into its more formal counterpart. All of these words have the same meaning, thus they are all turned into “good,” for example. As part of this study, a wordnet lemmatizer [31] was employed, which is an English lexical database that aims to construct organised semantic associations between words. Although it does not have a built-in lemmatization feature, this is one of the most widely used and oldest. In order to use the wordnet lemmatizers provided by NLTK, the following requirements must be followed.

---

### Algorithm

---

```

pip install nltk
import nltk
nltk.download('wordnet')
from nltk.stem import WordNetLemmatizer #In order to lemmatize,you need to create an instance of the
WordNetLemmatizer() and call the lemmatize() function on a single word.
lemmatizer = WordNetLemmatizer() # Init the Wordnet Lemmatizer
print(lemmatizer.lemmatize("bats")) # Lemmatize Single Word
#> bat
print(lemmatizer.lemmatize("are"))
#> are
print(lemmatizer.lemmatize("feet"))
#> foot

```

---

### 3.3 Feature Extraction

Feature extraction in data mining is a process that involves steps to reduce the amount of data available to describe large data sets. When analyzing the mood of a complex text, one of the main problems arises from the number of variables. Usually, to analyze complex and large text, large amounts of memory and processing power are required. This can make the classification algorithm more suitable for training samples and lead to poor generalizations with new samples. The researchers have discussed that in applications involving many attributes, feature extraction is similar to dimensionality reduction. Applying feature extraction techniques to the input data before passing the input data to the classification algorithm results in refining the classifier model accuracy.

This section introduces a new Element Extract technique for extracting elements from a given text. Using context words to represent or extract meaning from a big body of text is known as “counter vectorization.” Several mutual constraints are imposed to each word in order to determine the possible word correspondences. Text data can also be used to derive useful functions such as Term Frequency (TF) and Inverse Document Frequency (IDF). Eqs. (1) and (2) are used by TF and IDF to determine the frequency with which a phrase appears in a document.

$$\text{Term frequency} = \frac{\text{No. of times term } (t) \text{ appears in document}}{\text{Total no. of terms in a document}} \quad (1)$$

$$\text{Inverse document frequency} = \log \frac{\text{Total no. of documents}}{\text{No. of documents with term}(t)} \quad (2)$$

The TF-IDF technique is used to extract useful information from tweets. In this research work, 6288 features are extracted as document-term matrix TF-IDF features.

### 3.4 Dimensionality Reduction Using Latent Dirichlet Allocation

The LDA technique is used to lessen the sum of features after extracting the data. Latent topic models (LDAs) are probabilistic topic models, where each document is signified by a random mixture of latent themes. It is used to determine the latent topic structure from observable data by labelling each latent topic as a distribution over a predefined set of words. A two-step technique is usually used to generate the words in a document. In the first phase, each document is assigned a random distribution topic. In LDA, a word is a distinct piece of data from a vocabulary directory containing  $1, \dots, V$  words with a succession of  $N$  words in  $W$ .

During the production of a corpus, the LDA is observed for the three-layered representation and the parameters are investigated. The subject variables at the document level are examined in each document. In LDA, every word in the document is evaluated for word-level variables. The generating process of LDA is depicted as a joint distribution over a random variable. Eq. (3) is used to calculate the  $k$ -dimensional Dirichlet random variable's probability density function. Eq. (4) is used to evaluate successively the joint distribution of a topic mixture and the likelihood of a corpus (5).

$$p(\aleph|\pi) = \frac{\Gamma\left(\sum_{i=1}^k \pi_i\right)}{\prod_{i=1}^k \Gamma(\pi_i)} \aleph_1^{\pi_1-1} \dots \aleph_k^{\pi_k-1} \quad (3)$$

$$p(\aleph, x, y|\pi, \mu) = p(\aleph|\pi) \prod_{n=1}^N p(x_n|\aleph) p(y_n|x_n, \beta) \quad (4)$$

$$p(D|\pi, \mu) = \prod_{d=1}^M \int p(\aleph_d|\pi) \times \left( \prod_{n=1}^{N_d} \sum_{x_{dn}} p(x_{dn}|\aleph_d) p(y_{dn}|x_{dn}, \mu) \right) d\aleph_d \quad (5)$$

The Dirichlet parameter,  $M$ , the document,  $N$ , the number of words,  $x$ , the per-word topic assignment, and  $y$ , the observed word are all denoted by in this example.

Hidden variable posterior distribution calculation is critical for LDA model inferential tasks. Although the exact inference of the posterior distribution of the hidden variable is a difficult challenge. The RF is used to classify the data after the reduced features have been obtained.

### 3.5 Classification

The semantic orientation of the text decides whether a statement or a review is positive or negative for the specific subject matter. It is accomplished by ascertaining the semantic orientation of individual words. Two methods, sentimental classification and feature-based opinion mining, can be used to identify a product's or review's opinion. Opinion words are frequently referred to as sentimentality words. Positive opinion words reflect desired states, whereas negative opinion words represent unfavourable states. Wonderful, good, gorgeous, and astonishing are a few examples of positive opinion adjectives. Negative opinion terms include horrible, poor, bad, and worst. Opinion phrases and idioms, in addition to opinion words, play vital functions. They are referred to together as the opinion lexicon.

There are two types of data mining techniques: supervised and unsupervised. For the goal of data mining, each of these methods has its own set of algorithms. For the final sentimental categorization in

this study, a supervised learning method is applied. Supervised learning is a data mining method that infers an attribute using a function constructed from labelled training data. A series of training examples makes up the training data. In supervised learning, data classification is divided into two stages, namely the training and classification phases. A classifier is initially created to describe a predetermined set of data classifications or ideas. Finally, classification accuracy is determined by comparing the classified output to a small set of test data produced from training data.

### 3.6 Random Forest Algorithm

These algorithms represent different random elements to create different decision factories in sets. If there are classification problems, the results of these trees are summarised for the final forecast. When creating ensemble classifiers, randomization plays an important role in creating a wide range of models based on deterministic algorithms. Using integrated methods, several models are combined to improve the generalizability of the resulting classifiers. Traditionally, aggregation methods relied on deterministic algorithms with randomized process to generate various options.

Representatives of deterministic algorithms are Bagging, RF, Randomized C4.5 and Random Subspace. Individual DT and correlation among base trees are key issues that decide the RF classifier's performance. Because of this, the proposed enhanced random forest optimizes a large number of decision trees by selecting only uncorrelated data and good trees with high classification accuracies. The tree selection process has the following steps:

1. Identifying and selecting only the good trees with high classification accuracy.
2. The correlation is measured between the selected good trees.
3. Based on the measured correlation, only uncorrelated trees are selected.

RF algorithm is a popular technique for building machine learning systems. This is a supervised ML method proposed by Leo Breiman. RF is an integrated algorithm. A set consists of separately trained algorithms called core algorithms, whose predictions are combined to predict new events. RF uses the decision tree as the main algorithm. Generate multiple decision trees and combine the results of these decision trees as a final decision. RF introduces randomness in two ways:

1. Data is sampled at random to create bootstrap samples.
2. Random assortment of attributes or input features for generating individual base decision trees.

The graphical depiction of the random forest tree structure is represented in [Fig. 2](#).

---

#### Algorithm

---

##### Pseudocode of random forest

##### Input:

*Number of trees in random forest "n"*

*Number of feature vectors "M"*

*Training samples "N"*

*Proportion of feature vectors considered to build tree "m"*

##### Output:

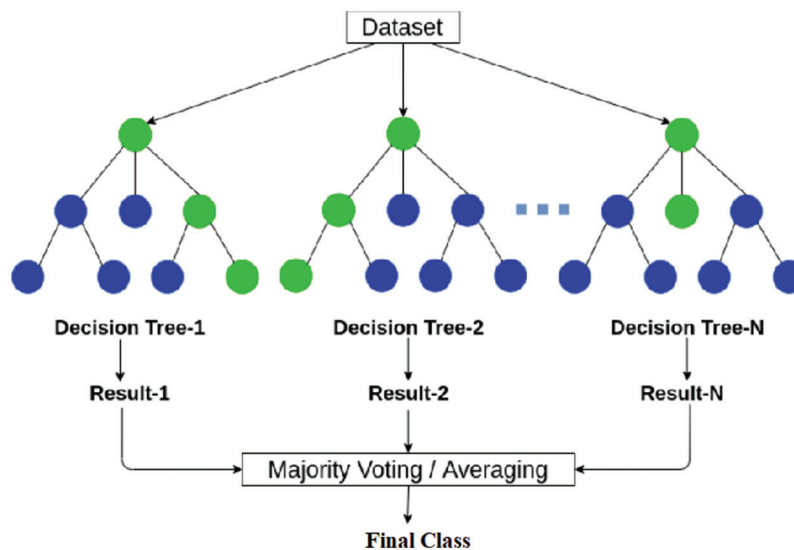
*Ensemble "m"*

1.  $E \leftarrow 0$
- 

(Continued)

**Algorithm (Continued)**

2. For  $i = 1$  to  $n$  do
3.  $N^i \rightarrow$  Bootstrap sample ( $N$ )
4.  $A^i \rightarrow$  Random select ( $m$ )
5.  $C^i \rightarrow$  Build random forest ( $N^i, A^i$ )
6.  $EU\{C^i\} \rightarrow E$
7. End for
8. Return  $E$
9. End algorithm



**Figure 2:** Graphical depiction of random forest classifier

When RF is used for cataloguing, the results of the basic decision trees are combined with a majority data to obtain better results.

## 4 Results and Discussion

This section provides detailed information on experimental results and discussions on proposed metrics, including performance metrics, experimental setup, dimensions, and benchmarking. The proposed system is realized using Python with 4 GB RAM, 1TB hard drive and a 3.0 GHz Intel i5 processor. The presentation of the system has been benchmarked against other classification methods. Finally, the performance parameter of the proposed scheme is evaluated for accuracy, recall, recall and F-measure.

### 4.1 Parameter Evaluation

The parameter performance factor is distinct as the regular measurement of results and outcomes to provide reliable information about the performance of the proposed system. In addition, performance measurement involves recording, storing, and analysing information about a group or individual's performance. Confusion metrics [31] for evaluating data classification are shown in Tab. 1. and can be understood in the following terms.



**Table 1:** Confusion matrix

	Predicted negative	Predicted positive
Actual negative	True negative (TN)	False negative (FN)
Actual positive	False positive (FP)	True positive (TP)

The mathematical equation of accuracy, precision, f-measure, and recall are signified in Eqs. (6)–(9).

$$Accuracy = \frac{TN + TP}{TP + TN + FN + FP} \times 100 \tag{6}$$

$$F - measure = \frac{2TP}{(2TP + FP + FN)} \times 100 \tag{7}$$

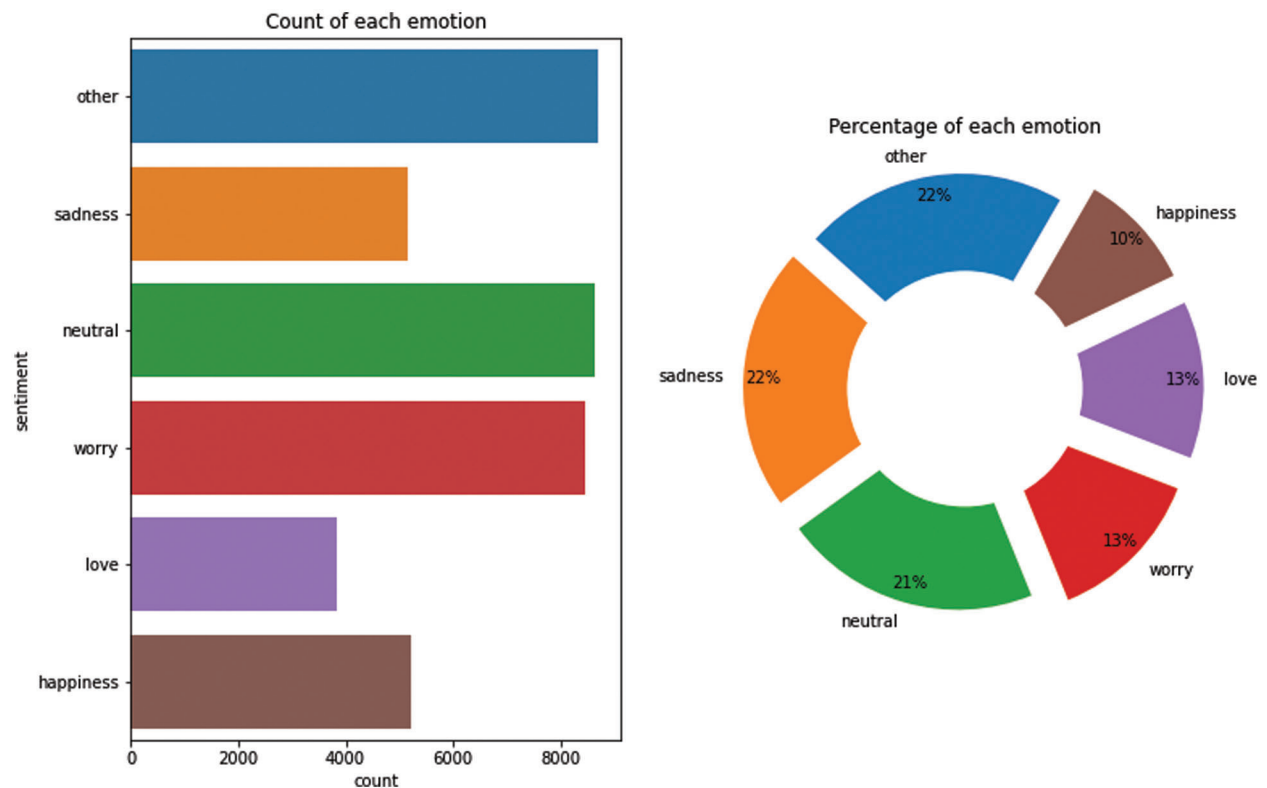
$$Precision = \frac{TP}{(FP + TP)} \times 100 \tag{8}$$

$$Recall = \frac{TP}{(FN + TP)} \times 100 \tag{9}$$

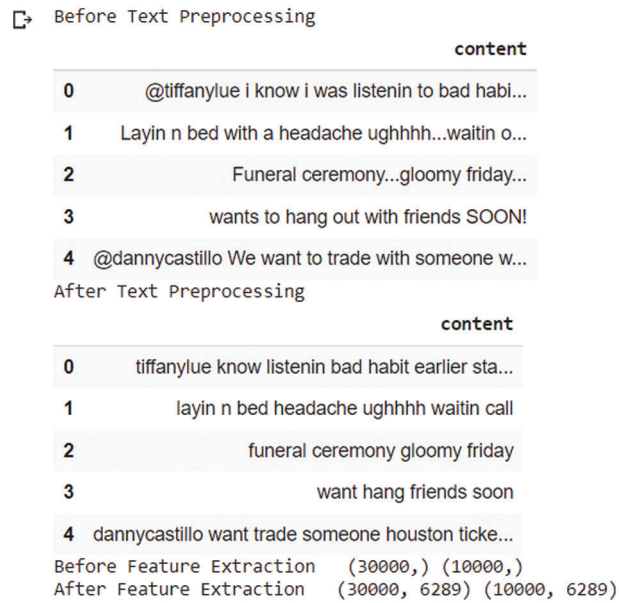
where TP is signified as true positive, TN is indicated as true negative, FP is specified as false positive, and FN is indicated as false negative.

**4.2 Performance Analysis of Proposed Pre-Processing (Wordnet Lemmatizer) Technique**

In this section, we will provide the graphical representation for imbalance dataset [32] in Fig. 3., where Fig. 4. shows the sample tweets before pre-processing and after applying pre-processing method.



**Figure 3:** Graphical representation of imbalance dataset before applying pre-processing technique

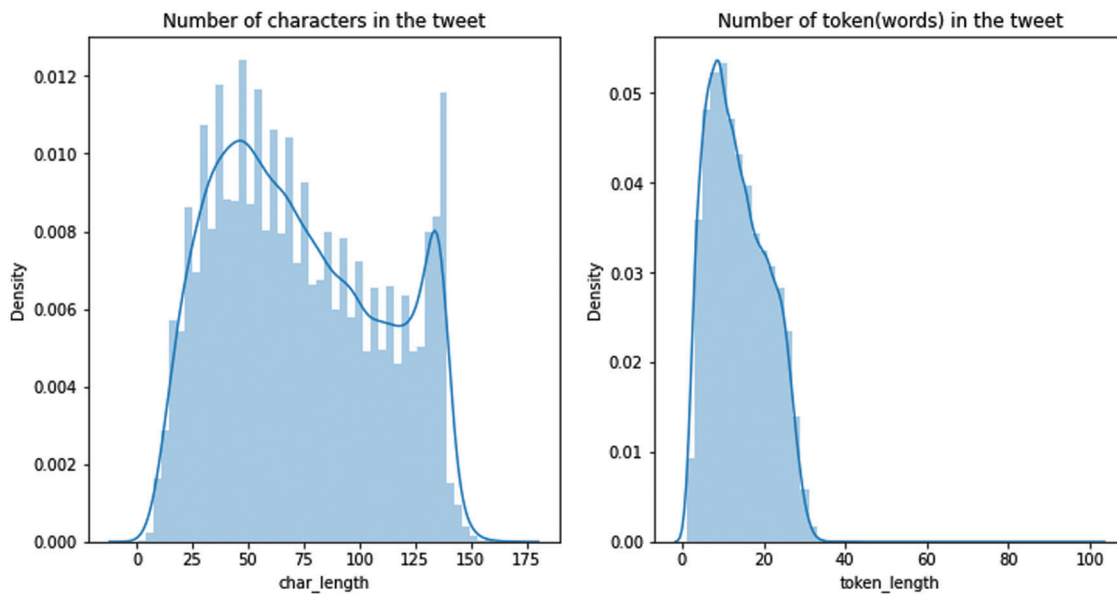


**Figure 4:** Sample tweets before applying and after applying pre-processing technique

In the above Fig. 4. the “@”, “!”, “...” are effectively removed by the proposed pre-processing technique. The unwanted words are also removed and it is clearly shown in the above Fig. 4.

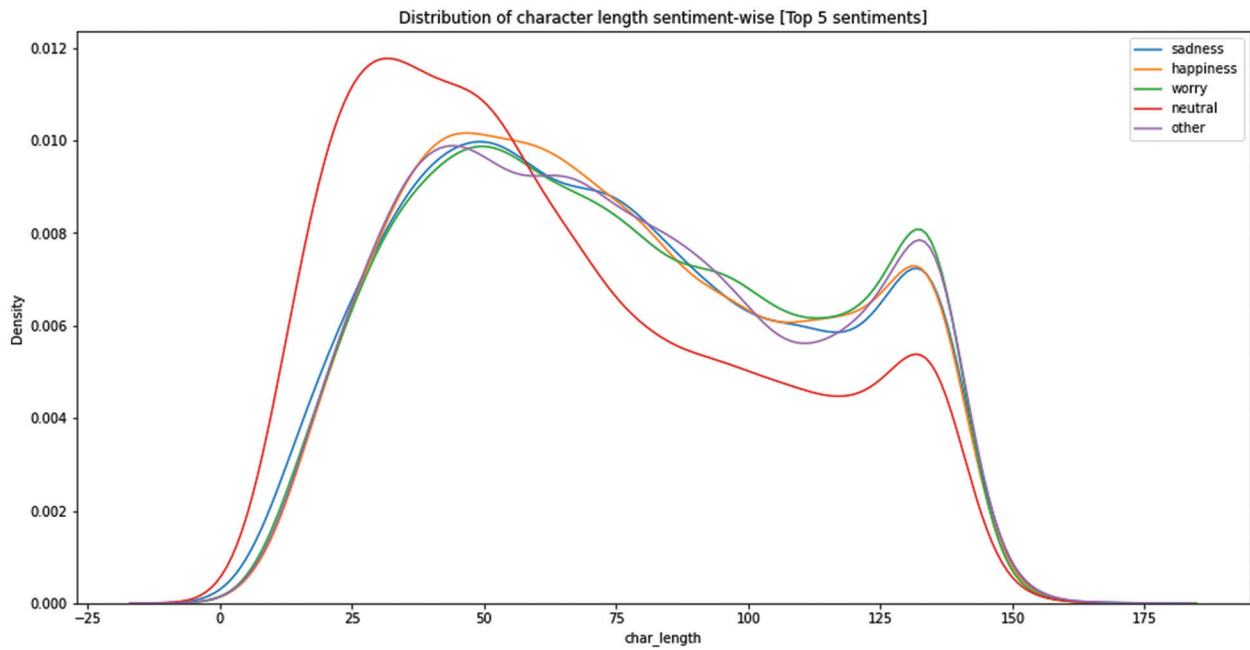
### 4.3 Validation of Character Length and Token Length

In this section, the graphical representation (Fig. 5.) shows the distribution of overall character length and token length.

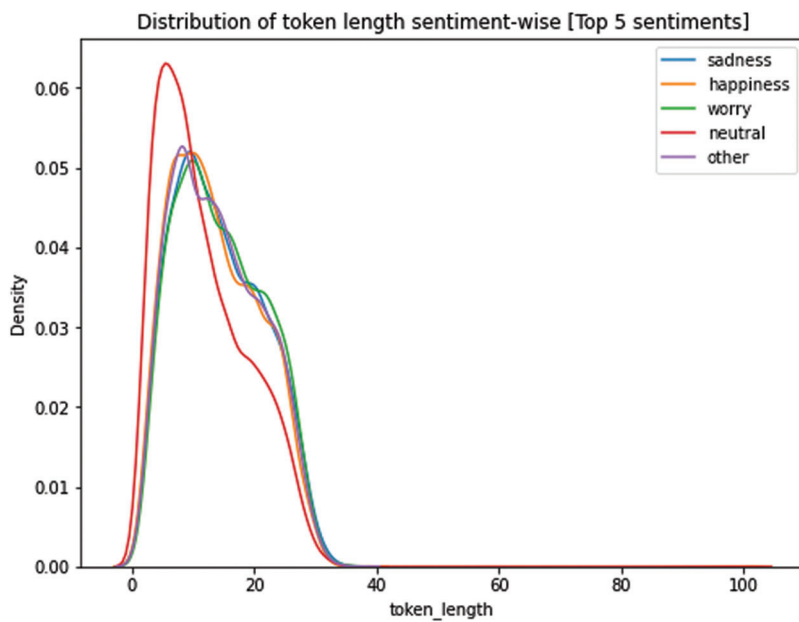


**Figure 5:** Overall distribution of character and token length

After applying the proposed method, top 5 sentiments are considered for distribution checklist. Here, multi-class text classifications includes worry, neutral, sadness, happiness and others are considered. Figs. 6. and 7 shows the graphical representation of distribution of character and token length wise for Top 5 sentiments.



**Figure 6:** Graphical representation of distribution of character lengthwise for top-5 sentiments



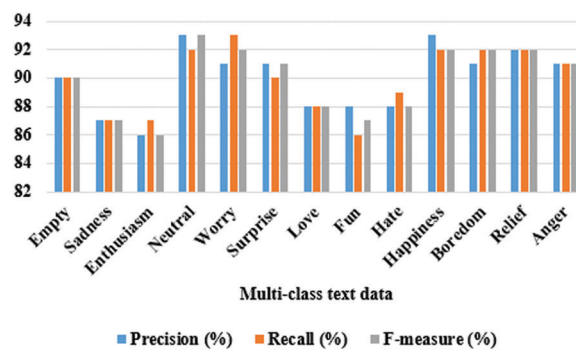
**Figure 7:** Graphical representation of distribution of token lengthwise for top-5 sentiments

#### 4.4 Performance Analysis of Proposed Random Forest Classifiers

In this section, the validation of RF is given in terms of recall and F-measure, where the overall classification accuracy is 90% for RF. In the datasets, there are 13 different types of emotions are labelled as multi-class and each class, the results shows the comparison of precision, recall and F-measure, which is exposed in [Tab. 2](#) and [Fig. 8](#).

**Table 2:** Experimental analysis of proposed RF for multi-class text emotion data

Multi-class texts emotions	Precision (%)	Recall (%)	F-measure (%)
Empty	90	90	90
Sadness	87	87	87
Enthusiasm	86	87	86
Neutral	93	92	93
Worry	91	93	92
Surprise	91	90	91
Love	88	88	88
Fun	88	86	87
Hate	88	89	88
Happiness	93	92	92
Boredom	91	92	92
Relief	92	92	92
Anger	91	91	91
Average	89.92	89.92	89.92



**Figure 8:** Graphical representation of proposed RF for multi-class text data in terms of precision, recall and F-measure

From the above [Fig. 8](#), it is clearly proves that RF achieved less performance in enthusiasm, love, fun and hate emotions, however RF achieved better performance in neutral and worry emotions. The precision, recall and F-measure are same for the empty, relief and anger emotions. In the precision experiments, the RF achieved high performance (i.e., 93%) on neutral and happiness emotions, where it achieved high recall value (i.e., 93%) only on worry emotions. Finally, in the f-measure experiments, the RF achieved 93% only on

neutral emotions. The average precision of RF is 89.92%, average recall is 89.92%, overall classification accuracy is 90% and average F-measure is 89.92%, which shows that RF achieved better performance in multi-class emotion text data. However, the overall classification accuracy is limited, since the dataset has more than five emotions for classification. This needs to be further enhanced by modifying the proposed model as future work.

## 5 Conclusion

Sentiment analysis at text is not always sufficient for analytical applications, as it does not specifically focus on opinions; In other words, emotions are not the purpose of existence. An element usually has many aspects, and people think about these aspects differently. For example, if the document represents the same entity, the document does not show that people have a positive opinion of all aspects of the item. Likewise, a negative opinion in a document does not mean that people are negative about all aspects of existence. Hence, careful analysis is needed to identify possible aspects and determine if a tweet is positive or negative. TSA is one of the new areas of research designed to analyze and identify users' feelings and attitudes. The suggested technique is divided into two stages: pre-processing and tweet categorization. Twitter's data is pre-processed by deleting extraneous smileys from tweets and accounting for lost value. TSA employs pre-processed data and the RF algorithm. The suggested method's experimental research is validated using emotion-related text data, demonstrating the superiority of the proposed methodology. The suggested approach on a multi-class emotion text dataset improves the classification rate of input data. In the future, we intend to investigate various methods of learning implicit and explicit associations in order to capture complicated data properties, which may aid in improving the performance of the suggested model.

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] G. Appel, L. Grewal, R. Hadi and T. Stephen, "The future of social media in marketing," *Journal of the Academy of Marketing Science*, vol. 48, no. 1, pp. 79–95, 2020.
- [2] L. Yue, W. Chen, X. Li and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, 2021.
- [3] A. Alayba, V. Palade, M. England and R. Iqbal, "Arabic language sentiment analysis on health services," *IEEE Access*, vol. 109, pp. 114–118, 2017.
- [4] B. Batrinca and P. C. Treleaven, "Social media analytics: A survey of techniques, tools and platforms," *AI Society Journal*, vol. 30, no. 1, pp. 89–116, 2015.
- [5] A. Giachanou and F. Crestani, "Like it or not: A survey of twitter sentiment analysis methods," *ACM Computer Surveying*, vol. 49, no. 2, pp. 752–771, 2016.
- [6] F. Cambria, L. Poria, R. Gelbukh and G. Thelwall, "Sentiment analysis is a big suitcase," *IEEE Intelligent Systems*, vol. 32, no. 6, pp. 74–80, 2010.
- [7] L. Chaturvedi, F. Cambria, A. Welsch and S. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges," *International Journal of Fusion*, vol. 44, no. 4, pp. 65–77, 2019.
- [8] F. Carvalho and A. Plastino, "On the evaluation and combination of state-of-the-art features in Twitter sentiment analysis," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1887–1936, 2021.
- [9] M. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*, vol. 55, no. 2, Cambridge: Cambridge University Press, pp. 109–131, 2020.
- [10] F. Cámara, L. Valdivia and R. Montejoraez, "Sentiment analysis in twitter," *Natural Language Engineering*, vol. 20, no. 1, pp. 1–28, 2019.

- [11] T. Mikolov, I. Sutskever, K. Chen and G. S. Corrado, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111–3119, 2018.
- [12] J. Pennington, R. Socher and N. Manning, "GLOVE: Global vectors for word representation," *EMNLP, Singapore*, vol. 11, pp. 189–211, 2014.
- [13] S. Manikandan, P. Dhanalakshmi, K. C. Rajeswari and A. Delphin Carolina Rani, "Deep sentiment learning for measuring similarity recommendations in twitter data," *Intelligent Automation & Soft Computing*, vol. 34, no. 1, pp. 183–192, 2022.
- [14] M. Xing, F. Cambria and X. Zhang, "Sentiment-aware volatility forecasting," *Knowledge Based System*, vol. 176, no. 1, pp. 68–76, 2021.
- [15] K. Li, L. Wu and F. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," *Information Process Management*, vol. 57, no. 5, pp. 102–212, 2021.
- [16] F. Xing, R. Cambria and M. Welsch, "Intelligent asset allocation via market sentiment views," *IEEE Computer Intelligent Magazine*, vol. 13, no. 4, pp. 25–34, 2020.
- [17] A. Araújo, L. Pereira and R. Benevenuto, "A comparative study of machine translation for multilingual sentence-level sentiment analysis," *Information Science Journal*, vol. 512, pp. 1078–1102, 2021.
- [18] F. Hussain and F. Cambria, "Semi-supervised learning for big social data analysis," *Neurocomputing*, vol. 275, pp. 1662–1673, 2021.
- [19] M. Emadi and M. Rahgozar, "Twitter sentiment analysis using fuzzy integral classifier fusion," *Journal of Information Science*, vol. 46, pp. 1–17, 2019.
- [20] Z. Jianqiang and G. Xiaolin, "Deep convolution neural networks for twitter sentiment analysis," *IEEE Access*, vol. 119, pp. 215–234, 2021.
- [21] P. Ray and A. Chakrabarti, "A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis," *Applied Computer Information*, vol. 119, pp. 219–237, 2019.
- [22] R. Speiser, F. Miller and L. Ipel, "A comparison of random forest variable selection methods for classification prediction modelling," *Expert Systems with Applications*, vol. 134, no. 10, pp. 93–101, 2021.
- [23] M. Qaiser and K. Ali, "Text mining: Use of TF-IDF to examine the relevance of words to documents," *International Journal of Computer Applications*, vol. 181, no. 1, pp. 25–29, 2021.
- [24] R. Kim and F. Gil, "Research paper classification systems based on TF-IDF and LDA schemes," *Human-Centric Computing and Information Sciences*, vol. 9, no. 1, pp. 1–21, 2020.
- [25] F. Henry and L. Twiscraper, "A collaborative project to enhance twitter data collection," in *14th ACM Int. Conf. on Web Search and Data Mining, India*, vol. 1, pp. 886–889, 2019.
- [26] M. Symeonidis, N. Effrosynidis and R. Arampatzis, "A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis," *Expert Systems with Applications*, vol. 110, no. 2, pp. 298–310, 2019.
- [27] K. Jianqiang and R. Xiaolion, "Comparison research on text pre-processing methods on twitter sentiment analysis," *IEEE Access*, vol. 51, pp. 2870–2879, 2017.
- [28] S. Manikandan and M. Chinnadurai, "Evaluation of students' performance in educational sciences and prediction of future development using tensorflow," *International Journal of Engineering Education*, vol. 36, no. 6, pp. 1783–1790, 2020.
- [29] S. Manikandan, P. Dhanalakshmi, R. Priya and K. Mary, "Intelligent and deep learning collaborative method for E-learning educational platform using tensorflow," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 10, pp. 182–198, 2021.
- [30] F. Haghghi, L. Jasemi, J. Hessabi and F. Zolanvari, "PyCM: multiclass confusion matrix library in python," *Journal of Open Source Software*, vol. 3, no. 25, pp. 729–741, 2019.
- [31] K. Komori and E. Eguchi, "Introduction to imbalanced data," in *Statistical Methods for Imbalanced Data in Ecological and Biological Studies*, Vol. 12. Tokyo: Springer, pp. 1–10, 2018.