ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2020.0302

REAL-TIME EMOTION RECOGNITION OF TWITTER POSTS USING A HYBRID APPROACH

Anjali Deshpande and Ratnamala Paswan
Department of Computer Engineering, Pune Institute of Computer Technology, India

Abstract

The analysis of social media posts is a challenging task, particularly the recognition of user emotions. Text is one of the most common mediums used by humans to express emotion, particularly on social media platforms. As emotions play a pivotal role in human interaction, the ability to recognize them by analyzing textual content has various applications in human-computer interaction (HCI) and natural language processing (NLP). Previous studies on emotion classification used bag-of-words classifiers or deep learning on static Twitter data. Our proposed model is a hybrid approach that uses a combination of keyword-based and learning-based models to perform textual emotion recognition on Twitter posts obtained in real-time. Textual feature extraction is carried out by standard Natural Language Processing (NLP) techniques such as Part-of-Speech (PoS) tagging and topic modeling along with classification done using the random forest algorithm. Results show that our proposed model performs better in comparison to the traditional Unison model with an average accuracy that approximates to 88.39%.

Keywords:

Emotion Recognition, Text Mining, Random Forest, Natural Language Processing, POS Tagging, Topic Modeling

1. INTRODUCTION

Emotions can be defined as the result of certain chemical changes in the nervous system that can be attributed to neurophysiological reactions associated with thoughts, feelings, behavioural responses or the degree of satisfaction or dissatisfaction. Human beings display a variety of emotions on a daily basis in the process of interacting to share their feelings. Such expressions of emotion are verbal and/or non-verbal, wherein the most common means of expressing emotion verbally is speech, while the non-verbal means are generally letters or gestures.

Emotions exhibited while sharing opinions, ideas, and feelings are conveyed through mediums that, in the present-day context, could be broadly classified as online and offline. In the offline aspect, emotions are conveyed in the form of letters or via oral communication. However, in the past few years, there has been a sudden increase in people displaying their thoughts through online media such as social media apps or websites. Social media platforms such as Facebook, Twitter, Instagram, etc. have gained momentum as standard platforms for people spanning across continents to express their emotions online. Nowadays, people have become more and more prone to expressing their opinions, daily life activities, feelings, etc. on such medium. Hence, the social media platforms have become huge repositories of emotion data that could be utilized effectively for research and analysis.

Emotion recognition plays a substantial role in facilitating Human-Computer Interaction (HCI), which is the key aspect that enables computers to understand and interpret human emotions. A wide variety of parameters have been used in the past by researchers for emotion recognition such as facial expressions, gestures, heart rate, blood pressure, EEG signals as well as text.

This paper strives to deliver an efficient textual emotion recognition system, as text is one of the most common mediums used by humans to express emotion. Explicit emotion recognition, which analyses hidden emotion within the text, is an important Natural Language Processing (NLP) task. It requires semantic interpretation as opposed to implicit emotion recognition that doesn't take into account contextual information. It has become an attractive topic for researchers as it has a variety of applications in the form of e-learning, data mining, recommender systems, psychology, information learning systems, business, and education as it facilitates the understanding and interpretation of human emotions. Also, text forms the basis for HCI via chatbots, online forums, SMS, emails, etc. and most importantly online social media platforms like Facebook, Twitter, etc.

1.1 HISTORY OF EMOTION RECOGNITION

The first study of emotions was carried out by Darwin via the analysis of facial expressions and body language of humans as well as animals [1]. Paul Ekman defined six basic emotion categories such as anger, disgust, fear, joy, sadness and surprise [2]. This model is one of the most popular ones for building emotion models as it is based on collective information from previous studies.

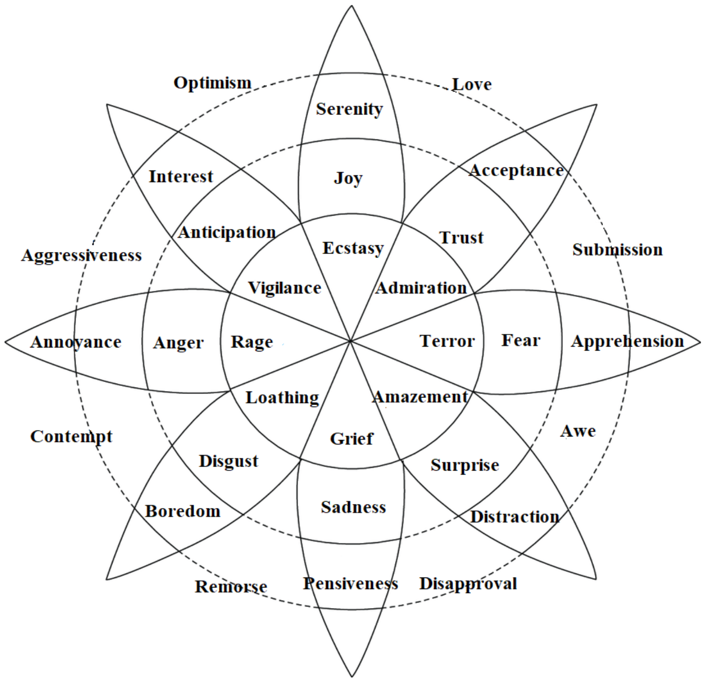

Fig.1. Robert Plutchik's wheel of emotions [3]

The supplementary emotions of trust and anticipation were later on added by Robert Plutchik [3] who presented his wheel of emotions as shown in Fig.1. Here, the emotions that are similar occur close to each other whereas opposite emotions are placed 180^0 far. This model is based on a total of eight emotions, each with 3 levels of intensity as depicted in Fig.1. An example for the 3 levels of intensity for one such emotion are ecstasy (high level), joy (medium level), and serenity (low level). However, this study has focused only on the emotions and not their intensities.

Based on that, a Profile-of-mood-states (POMS) was developed which is basically an instrument that defines a six-dimensional emotional-state representation; each dimension representing one of the six basic emotions. This instrument can be extended for the detection of more emotional categories. Our approach defines a twelve-dimensional POMS that makes use of a combination of basic as well as supplementary emotion categories such as anger, trust, anticipation, disgust, fear, joy, depression, surprise, fatigue, vigour, tension and confusion. Each dimension represents one of the formerly mentioned twelve emotion categories. We have included the supplementary emotions of depression, fatigue, tension and confusion as these are crucial emotions for the assessment of the mental state of any individual.

1.2 EMOTION RECOGNITION APPROACHES

Emotion recognition is mostly done on two levels: low-level/coarse-grained level analysis or high-level/fine-grained level analysis. The low-level analysis encompasses a binary classification of the text into positive or negative which is also called as Sentiment Analysis. The high-level analysis is basically a higher form of sentiment analysis which is nothing but a further classification into crisp emotional categories. Basically, there are four different approaches to emotion recognition: machine learning based, keyword-based, deep learning based and hybrid/combination approach as explained in Table.1.

Table.1. Emotion Recognition Approaches

No.	Approach	Characteristics
1.	Keyword based	• Traditional and Easy • Emotion recognition via predefined emotion words (keywords) which are detected via some rules and vocabularies. • E.g. WordNet dictionary
2.	Machine learning based	• Builds classification model using ML algorithms. • Training with large emotion data and emotion recognition of new incoming data. • E.g. SVM, Naïve Bayes, Decision tree, Random forest, etc.
3.	Deep learning based	• Emotion recognition via deep learning models, no explicit feature extraction required and needs large amount of training data e.g. Convolutional Neural Network (CNN), etc.
4.	Hybrid/ Combination	• Uses a combination of the above approaches.

Emotion recognition is undoubtedly an important research aspect in the field of text mining. Timely detection of the stress-state or suicidal tendency of a person on the basis of emotion recognition is one of the examples why the research is an indispensable necessity. Also, the public mood of a current topic can be inferred effectively using real-time data such as electoral tweets, trending topics, etc. Also, the analysis on real-time data is crucial as it proves quite beneficial for deployment in real-time applications.

1.3 RANDOM FOREST CLASSIFIER

Random Forest (RF) is a very popular machine learning algorithm mainly used for classification or regression purposes. It can be mostly thought of as a collection of individual decision trees. It mostly works on the concept of ensemble learning. Ensemble methods are those that use multiple executions of a task or multiple algorithms to achieve better predictive performance as compared to that obtained with respect to single execution or single algorithm.

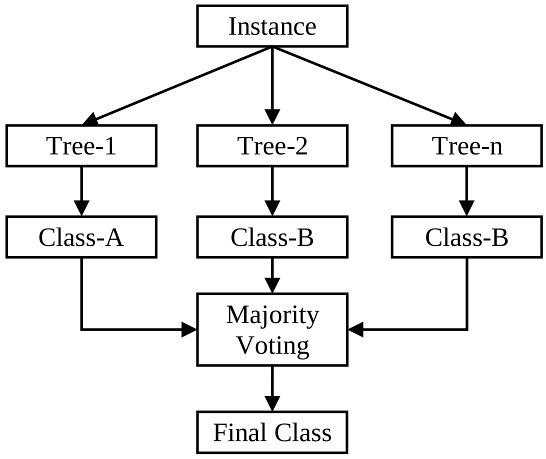

Fig.2. Random Forest Classifier [4]

With RF, each decision tree predicts a class, called as vote, for the test sample individually. The final prediction of RF is the one with maximum votes (in case of classification problem) or the value obtained after averaging the votes (in case of regression problem). The working of RF algorithm for classification is as depicted by Fig.2. The advantage of using RF is that it reduces over-fitting of the predictions to training set.

1.4 CONTRIBUTIONS

The previous work done in the area of emotion recognition had the following limitations:

- Emotion recognition was carried out only on static i.e. offline dataset
- Either of the single approaches were used for classification.
- No single model for predicting multiple emotion classifications.

The contributions of our work are as follows:

- Increasing the overall accuracy of emotion recognition via machine learning algorithms assisted by feature extraction using NLP algorithms.
- Analysis on Twitter posts obtained in real-time.

- Classification by modelling a twelve-dimensional mood-state representation.
- Analysis is done on both levels, low-level (sentiment analysis) as well as high-level (crisp emotion recognition) for twelve emotion categories formerly mentioned.
- The proposed model is a hybrid model, that is, a combination of keyword-based and machine learning based approaches.

2. LITERATURE SURVEY

From a long time, researchers have shown interest in the study of human emotions. Darwin [1] was one of the first to study emotions of both humans as well as animals. The six basic human emotions were explained by Ekman [2] which were anger, disgust, fear, joy, sadness and surprize. The addition of trust and anticipation was later on done by Plutchik [3]. Textual emotion recognition is a wide area of research that is quite similar to sentiment analysis. The previous researches done in this field indicated that the analysis on emotion recognition is of two types viz. sentiment analysis and emotion analysis/affective computing and a clear differentiation between these two concepts is presented in [5].

Various approaches were used by researchers for the recognition of emotions as explained in section 1.2. A detailed comparative study of various emotion recognition approaches has been carried out in [14]. The methods in [7] [8] [11] - [13] [16] [18] and [19] use machine learning approach. The deep learning approach is harnessed in [6] [9] [10] and [15]. Additionally, [17] comparatively analyses both deep learning and machine learning approaches. The Table.2 gives a detailed overview of this research carried out in the field of emotion recognition [20] – [25].

Table.2. Literature survey

No.	Techniques	Inferences
[6]	**Methodology** • Use hashtags to create three large datasets labelled with the emotions expressed. • Propose a unified model for multiple emotion recognition **Algorithm used** • CNN	**Advantages** • Development of single classification model • Detection of eight emotions • Improved performance using the new Hashtag emotion corpus. **Limitations/Future Scope** • No semantic analysis done • Only static data used for analysis • Only hashtags used for labelling
[7]	**Methodology** • Used emotion-word hashtags for manual labelling of tweets. • Generated the first large corpus of word–emotion associations. • Mainly worked with six basic emotions. **Algorithm used** • Six one-vs.-all classifiers using SVM for each basic emotion.	**Advantages** • Improvement in classification performance. • Automatic detection of personality from tweets **Limitations/Future Scope** • The synonyms of the emotion words not considered. • Emotions of trust and anticipation were not considered
[8]	**Methodology** • Used social media as a resource for personality detection • Using personality predictive features, a novel corpus of 1.2M English tweets annotated with Myers-Briggs personality type and gender was generated. **Algorithm used** • Logistic regression	**Advantages** • Personality distinctions viz. introvert–extrovert (I–E) and thinking–feeling (T–F), predicted with high reliability **Limitations/Future Scope** • Other personality traits couldn't be distinguished
[9]	**Methodology** • Propose a multi-task DNN for representation learning, combining semantic classification and semantic information retrieval tasks. • Model maps arbitrary text queries and documents into semantic vector representation. **Algorithm used** • MT-DNN	**Advantages** • The MT-DNN robustly outperforms strong baselines across all web search and query classification tasks. **Limitations/Future Scope** • Application of this method to various other knowledge sources.
[10]	**Methodology** • Explored an application of deep recurrent neural networks to the task of sentence-level opinion expression extraction.	**Advantages** • RNNs outperformed previous traditional methods **Limitations/Future Scope** • Only considered word vectors as the features. • Pre-training not done

	• Evaluated deep RNNs against conventional, shallow RNNs that have only a single hidden layer. **Algorithm used** • RNN	
[11]	**Methodology** • Analyse electoral tweets for sentiment, emotion, purpose or intent behind the tweet • Two-step process: annotating text for sentiment, emotion, style, and categories and automatic classification. **Algorithm used** • SVM with 10-fold cross-validation	**Advantages** • Automatically classified tweets into emotional categories • Also developed supervised classifiers for detecting emotional state, the purpose as well as the stimulus behind the users' tweets. **Limitations/Future Scope** • Mostly handled the emotions pertaining to disgust or trust but not others • Past tweets were not considered
[12]	**Methodology** • Twitter has been hazardously plagued by bots • Propose a generalized machine learning model that can detect hazardous Twitter bots with high accuracy. **Algorithm used** • Random forest	**Advantages** • Achieved very high accuracy of 90.25% • Create a generalized model that can be deployed for accurate use directly after training **Limitations/Future Scope** • Model works well with Twitter data.
[13]	**Methodology** • Sentiment Analysis of Kannada documents • Propose an ensemble of classifiers with random forest technique to identify and test the polarity of the sentiment • Handle multi-class labels, identify sentiment polarity of comparative and conditional statements. **Algorithm used** • Random forest	**Advantages** • Improve the overall performance of sentiment analysis of Kannada documents done previously • Overall accuracy is improved from 65% to 72% indicating the efficiency of the proposed model. **Limitations/Future Scope** • Future scope emphasizes work with larger dataset and NLP techniques.

[14]	**Methodology**
	• Carried out a survey pertaining to various methods used for emotion recognition.
	• Identification of informative keywords in text is very important.
	• Four fundamental methods described are:
	  • Rule-based,
	  • Classical learning
	  • Deep learning
	  • Hybrid
		Advantages
		• Learning based methods perform better than other approaches
		• Text with traditional word distribution is favourable.
		• Also state the importance of using NLP techniques such as PoS-tagging.
		Limitations/Future Scope
		• State the following challenges for emotion recognition
		  • Recognize the emotions implicitly expressed
		  • Data collection and language of concern
		  • Emphasized the use of real-time data for analysis in future.

[15]	**Methodology** • Propose an emotion embedding model that is created by developing a word embedding layer during the emotion classification phase. • Used corresponding hashtags of Tweets as emotion labels. • Classification into eight emotions based on Plutchik's model **Algorithm used** • CNN	**Advantages** • Model gave promising results • Emotion of "joy" predicted with high accuracy (0.73) as opposed to "Anger" (0.36). **Limitations/Future Scope** • Doesn't take into account the semantics of the text. • Doesn't work on real-time data.
[16]	**Methodology** • Carried out emotion recognition via four supervised learning classifiers. • Only six basic emotions classified. • ISEAR (International Survey on Emotion Antecedents and Reactions) dataset used. **Algorithm used**	**Advantages** • Multinomial Naïve Bayes classifier outperformed others with an accuracy of 64.08%. • Also built a GUI for emotion prediction. **Limitations/Future Scope** • Model can be made more refined and precise. • No semantic analysis done. • Model mainly worked on static dataset

	• K-nearest neighbour, Decision tree, Multinomial Naïve Bayes and Support Vector Machine	
[17]	**Methodology** • Built a standard manually labelled Vietnamese social media corpus called UIT-VSMEC. • Six basic emotions taken into consideration. • Both Machine learning and deep learning approaches were used. **Algorithm used** • Random Forest, SVM, LSTM and CNN	**Advantages** • From Machine learning approach, SVM performs better • From Deep learning approach, CNN performs better. **Limitations/Future Scope** • Relative shortage of comments expressing anger, fear and surprise. • Real-time analysis not done.
[18]	**Methodology** • Emotion recognition of Bengali text. • Six basic emotions were recognized • Emotion corpus is built via collection from heterogenous sources and manual annotation technique. **Algorithm used** • SVM and Naïve Bayes	**Advantages** • SVM with an accuracy of 73% outperformed Naïve Bayes with 60% **Limitations/Future Scope** • Corpus size and no inclusion of real-time or semantic analysis.
[19]	**Methodology** • Made use of supervised classifiers for emotion recognition • Classification into four categories: Happy-Active, Happy -Inactive, Unhappy-Active and Unhappy -Inactive. • Labelled emotion corpus created using Russell's Circumplex model via a rule-based approach. **Algorithm used** • Eager learner: Naïve Bayes and Lazy learner: KNN	**Advantages** • Naïve Bayes, with an accuracy of 72.60%, performed better than KNN with 55.50%. **Limitations/Future Scope** • Working with other models as well as combination models.

3. PROPOSED METHODOLOGY

3.1 PROPOSED SYSTEM ARCHITECTURE

The architecture of our proposed model is described in Fig.3 and subsequently explains the steps of emotion recognition.

Step 1: Real-time data collection

The posts on Twitter aka tweets are obtained in real-time by creating an app via the Twitter API. For accessing the real-time tweets, we use the open-source java library Twitter4J. Via Twitter4J, a user can do the following tasks:

- Post tweets
- View user timeline
- Access real-time latest tweets
- Send and receive messages

Along with these tasks, this library also ensures the security and privacy of the user. The obtained tweets are stored in the database. The advantage of using real-time data is that analysis is done using the latest data whenever required.

Step 2: Data pre-processing

The raw data obtained from Twitter often contains special characters (e.g. #,?,!, etc.), emojis. Hence, those are removed.

1. *Stopword Removal*: Stopwords are those words that occur frequently and have no contribution in the actual context of a document (e.g. a, and, the, etc.). Hence, stopword removal is carried out.
2. *Stemming:* Also, most sentences contain the derivatives of a base word called as stem. Hence, reduction of the derivatives of words to their root form is carried out also called as stemming (e.g. "joyful", "joyous" reduced to the stem "joy"). Porter's stemming algorithm is used for this purpose.
3. *Tokenization:* The sentence is further divided into individual words called as tokens. This process is called as tokenization. All these steps are carried out to clean the data and prepare it for further analysis.

Step 3: Sentiment Analysis

It is a process of classifying a given text as expressing a positive, negative or neutral sentiment. For that purpose, we are using the SentiWordNet dictionary in which each word is given a score according to positive or negative information displayed. Using that scores, a tweet is probabilistically classified as positive, negative or neutral.

Thus, we achieve low/coarse-grained level analysis using above approach.

Step 4: Feature extraction using NLP

In this step, repetitive tweets' content, that occur mostly due to retweets by other users to a particular user's tweet, are removed using n-gram generation by using tweet-id. For extracting textual features to train our classification model, we used two NLP techniques viz. PoS-Tagging and Topic Modeling.

Part-Of-Speech tagging is also called as grammatical tagging in which the words of a particular text are tagged with the corresponding Part-of-Speech. This process is very helpful for identifying the most useful features which in turn are useful for identifying the emotions.

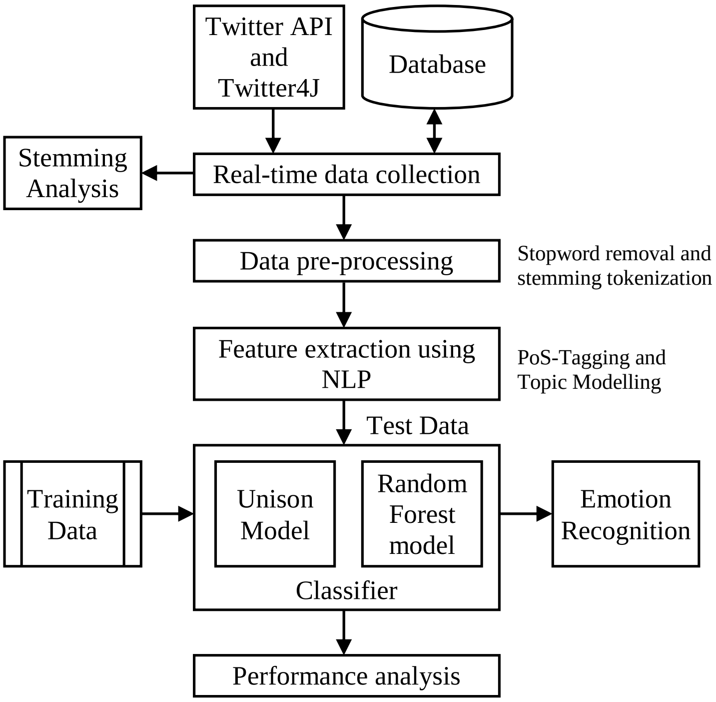

Fig.3. Proposed System Architecture

The emotion words mostly occur as nouns (basic tag as NN), adjectives (basic tag as JJ) or verbs (basic tag as VB) in a sentence and hence are extracted. The maximum entropy classifier or MaxEntTagger class available in stanford.nlp.tagger package in Java is used for this purpose. After that, the n-grams of tweets are generated. Then by using a text clustering algorithm, a similarity score is assigned to the tweets and similar tweets are grouped and placed in the same cluster.

Topic modeling is a text mining tool used to probabilistically find the abstract topics on which a document or a collection of text is based. Statistical algorithms are applied for discovering the latent semantic topics depicted by a textual collection. Topics are those words that are found to be most informative regarding that tweet. This approach is mostly used to distinguish different occurrences of the same word.

The significance of this technique is that the obtained features ensure a more generalized classifier output. Latent Drichlet Allocation (LDA) algorithm is used for this purpose. This algorithm is systematically explained in section 3.2.1 of this paper. This is a keyword-based approach in which the keywords or features are extracted for further analysis.

Step 5: Emotion classification

In this step, two classification models are built and trained with emotion-word features obtained from [6]. The first is the Unison model which is basically the traditional emotion classification model based on Bag-of-Words model. POMS is used for its implementation. The second model is our proposed Random Forest (RF) model. It is built by training it with emotion-labelled tweets. The random forest algorithm is used for solving Quadratic Programming (QP) problems and hence it is chosen to build a multiclass classification model for emotion detection.

The random forest algorithm is further explained in section 3.2.2 of this paper. The features obtained from the previous step are used as test data and, finally, the emotions of the tweets are detected and are displayed to the user. This is the machine learning-based approach with which we obtain high/fine-grained level analysis.

Step 6: Performance Analysis

The performance of the proposed novel method is analysed with respect to the traditional unison model. This is done using the performance measures such as accuracy, precision, recall and f1-score. A comparison of the performance of both the models is done graphically.

3.2 ALGORITHMS USED

3.2.1 Latent Dirichlet Allocation (LDA) Algorithm:

LDA describes how the documents in a dataset were created using a generative version. Consider that a dataset is a collection of D documents which, in turn, can be considered as a collection of words. Hence, this model indicates how each document acquires its words. Initially, let's assume that K topic distributions for our dataset, meaning K multinomials each containing Y elements, where Y is the number of terms in the corpus. Let y_i represent the multinomial for the i^{th} topic. The size of y_i is Y:$|y_i| = Y$. On the basis of these, the generative process of LDA can be modelled as:

Steps:

a) For each document:
   - i) Select a K-sized random distribution
   - ii) For each word occurrence:
     - (1) Select one topic from the K topics obtained via the distribution in (a) in a probabilistic manner, call it y_j
     - (2) Select one out of the Y words from y_j in a probabilistic manner.
   - iii) End

b) End

3.2.2 Random Forest:

The algorithm has two parts:

- **Forest generation**

Step 1: Suppose the total number of samples is N and the total number of attributes are M. The total number of attributes required to determine the splitting criterion for the classifier is m such that $m << M$.

Step 2: Bootstrap with replacement n times to generate a training sample from the N samples available. Predict the class of remaining samples and use them for error estimation. Choose m attributes and decide the splitting criterion at each node. Based on these, calculate the best split in the training set of samples. Grow each tree fully and make sure that no pruning is involved on any level.

- **Output prediction:**

For any new incoming sample xi,

Step 1: x_i is pushed down the currently generated tree until it achieves the position of a terminal node of the training samples. After that, xi is assigned the label X to which that node belongs. This process is repeated for each tree generated while bootstrapping with replacement, say n times.

Step 2: In case of classification, let $L_0, L_1, L_2,\ldots,L_m$ are the actual classes/labels (in case of multiclass classification with m classes). Thus, the final prediction y is given as

$$y = \max[n(L_0), n(L_1), n(L_2),..,n(L_n)] \tag{1}$$

Eq.(1) is for the prediction of classification output which is the same as the class that has the maximum number of predictions ($n(L_i)$ is the total number of predictions for class i).

In case of regression, let $y_1, y_2, y_3,\ldots,y_n$ be the outputs generated by n decision trees in the ensemble each. The final output is given as,

$$y = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2}$$

Eq.(2) is for predicting regression output which is the average of n outputs generated by each tree in the ensemble. Thus, the test sample x_i is classified into class y using Random Forest algorithm.

4. RESULTS AND DISCUSSIONS

An emotion recognition web application is developed using Eclipse IDE, MySQL and Apache Tomcat server. The coding is done in Java. Java servlets are created. The Twitter API is used to obtain real-time tweets for analysis by developers via the Twitter4J library. The retrieved tweets are preprocessed in suitable format. The SentiWordNet dictionary is used for sentiment analysis of the tweets. NLP algorithms are then used for extracting informative features. The Unison model is the previously developed model for emotion recognition based on POMS [6]. Proposed model for emotion recognition is developed as explained in section 3.1. The classification is performed using both the classifiers. The labels of real-time test data are assigned according to the probability of occurrence of emotion words predefined in [6] which later helps in performance evaluation of the classifiers. Their performances are expressed in terms of performance measures. The metrics of classification performance evaluation used are accuracy, precision, recall and f1-score.

$$\text{Accuracy} = \frac{\text{Correctly Predicted Observation}}{\text{Total number of Observation}} \tag{3}$$

$$\text{Precision} = \frac{\text{Correctly Predicted Positive Observation}}{\text{Total Predicted Positive Observation}} \tag{4}$$

$$\text{Recall} = \frac{\text{Correctly Predicted Positive Observation}}{\text{Total Positive Observation}} \tag{5}$$

$$\text{f1-score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \tag{6}$$

Accuracy is the measure of how perfectly the classifier classifies the incoming unknown data and is defined as the ratio of correctly predicted data samples to total data samples as shown in Eq.(3).

Precision is the fraction of predicted positive instances that are actually positive and is calculated as the fraction of correctly predicted positive observations to the total predicted positive observations (refer Eq.(4)).

Recall is the fraction of the actual positive instances classified as positive and it is the ratio of the correctly predicted positive observations to the total positive observations shown in Eq.(5).

Lastly, Eq.(6) defines f1-score which is said to be the harmonic mean of precision and recall. All these measures are used to evaluate the performance of both the classification models. These measures are as given in Table.3. As the data is real-time, the data is taken as the average of 10 executions.

Table.3. Performance measures: Unison Model vs. Proposed Random Forest Model

Metrics	Unison Model (%)	Random Forest (%)
Precision	79.45	83.70
Recall	78.41	77.64
F-Measure	79.11	81.31
Accuracy	82.74	88.39

From the Table.3, we can draw a comparative analysis between both these classification models. We can clearly see that the proposed random forest model with an accuracy of 88.39% outperforms the traditional unison model having 82.74% accuracy. The other performance measures also indicate the same except for recall of random forest classifier, which is comparable with that of the recall of unison model.

The classification performance of Unison model is compared with the proposed model and is displayed graphically in Fig.4 and Fig.5. The Fig.4 shows the total number of tweets that are classified by both the classifiers into each of the twelve emotion categories. For e.g. in case of the emotion category "joy", the unison model classifies 78 tweets into this category whereas random forest model classifies 90 tweets as such.

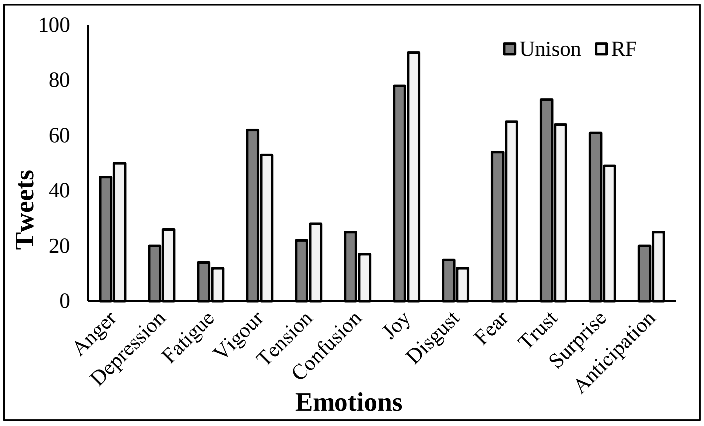

Fig.4. Total number of tweets predicted in each emotion category - Comparative Analysis

We can see also that, in most of the cases, the tweets misclassified by unison model are classified into the correct category by the proposed hybrid model. For e.g. the tweet ".... not that I am quite amazed by this …" is classified in the category "surprise" by unison model but random forest classifies it into the category "anticipation". The Fig.5 gives a graphical comparative analysis of both the classifiers in terms of their performance measures. All the values are in terms of percentage.

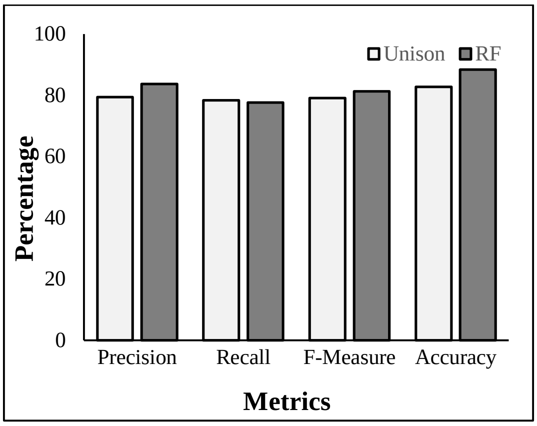

Fig.5. Performance measures - Comparative Analysis

Thus, we draw the following inferences from our experiments:

- The results obtained show that the proposed model performs significantly better than Unison model and contributes to the increase in the overall accuracy of emotion recognition.
- The proposed model successfully classifies 12 emotion categories.

The improved performance of the proposed model can be attributed to the inclusion of NLP algorithms and the hybrid approach employed for emotion recognition.

5. CONCLUSION

Emotion recognition is a vast area of research which has found its application in crucial fields such as detecting psychological disorders in individuals like anxiety or depression, measuring the public mood of a community (e.g. electoral tweets, etc.) or sentiment analysis. Our work proposes a novel hybrid method for emotion recognition using machine learning and NLP semantic analysis with a key focus on twelve emotions. For substantiating the proposed method, it is tested using tweets that are obtained from Twitter in real-time. The results show that our proposed model performs significantly better than the previous Unison model, and that it accomplishes the task in a comparatively shorter duration. This betterment is achieved by the inclusion of feature extraction using NLP techniques and classification using random forest algorithm.

The scope for applying the proposed method extends from vital analytics like the automatic real-time online user feedback analysis, user reputation monitoring, and so on. Future research work on this model could progress towards the inclusion of images and/or videos for assisting in emotion recognition, working with other social media data, and even considering other languages for emotion recognition analysis.

REFERENCES

[1] C. Darwin, "*The Expression of the Emotions in Man and Animals*", John Murray, 1872.

[2] P. Ekman, "An Argument for Basic Emotions", *Cognition and Emotion*, Vol. 6, No. 3-4, pp. 169-200, 1992.

[3] R. Plutchik, "*Emotion: Theory, Research, and Experience*", Academic Press, 1980.

[4] Random Forest Classification, Available at: https://towardsdatascience.com/random-forest-classification-and-its-implementation-d5d840dbead0

[5] M.D. Munezero, C.S. Montero, E. Sutinen and J. Pajunen, "Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text", *IEEE Transactions on Affective Computing*, Vol. 5, No. 2, pp. 101-111, 2014.

[6] N. Colneric and J. Demsar, "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model", *IEEE Transactions on Affective Computing (Early Access)*, pp. 1, 2018.

[7] S.M. Mohammad and S. Kiritchenko, "Using Hashtags to Capture Fine Emotion Categories from Tweets", *Computational Intelligence*, Vol. 31, No. 2, pp. 301-326, 2015.

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

[9] X. Liu, J. Gao, X. He, L. Deng, K. Duh and Y.Y. Wang, "Representation Learning using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval", *Proceedings of International Conference on North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 912-921, 2015.

[10] O. Irsoy and C. Cardie, "Opinion Mining with Deep Recurrent Neural Networks", *Proceedings of International Conference on Empirical Methods in Natural Language Processing*, pp. 720-728, 2014.

[11] S.M. Mohammad, X. Zhu, S. Kiritchenko and J. Martin, "Sentiment, Emotion, Purpose, and Style in Electoral Tweets", *Information Processing and Management*, Vol. 51, No. 4, pp.480-499, 2015.

[12] J. Schnebly and S. Sengupta, "Random Forest Twitter Bot Classifier", *Proceedings of 9th Annual Workshop and Conference on Computing and Communication*, pp. 506-512, 2019.

[13] Y. Hegde and S.K. Padma, "Sentiment Analysis using Random Forest Ensemble for Mobile Product Reviews in Kannada", *Proceedings of IEEE International Conference on Advance Computing*, pp. 777-782, 2017.

[14] N. Alswaidan and Mohamed El Bachir Menai, "A Survey of State-of-The-Art Approaches for Emotion Recognition in Text", *Knowledge and Information Systems*, Vol. 62, pp. 2937-2987, 2020.

[15] S. Park, B. Bae and Y. Cheong, "Emotion Recognition from Text Stories Using an Emotion Embedding Model", *Proceedings of IEEE International Conference on Big Data and Smart Computing*, pp. 579-583, 2020.

[16] Ahmad Fakhri Ab. Nasir, Eng Seok Nee, Chun Sern Choong, Ahmad Shahrizan, Abdul Ghani, "Text-Based Emotion Prediction System using Machine Learning Approach", *Proceedings of IOP Conference Series: Materials Science and Engineering*, pp. 1-13, 2020.

[17] Vong Anh Ho, Duong Huynh Cong Nguyen, Danh Hoang Nguyen, Linh Thi Van Pham, Duc Vu Nguyen, Kiet Van

Nguyen and Ngan Luu Thuy Nguyen, "Emotion Recognition for Vietnamese Social Media Text", *Proceedings of International Conference of the Pacific Association for Computational Linguistics*, pp. 319-333, 2019.

[18] H.A. Ruposh and M.M. Hoque, "A Computational Approach of Recognizing Emotion from Bengali Texts", *Proceedings of 5th International Conference on Advances in Electrical Engineering*, pp. 1-7, 2019.

[19] M. Suhasini and B. Srinivasu, "Emotion Detection Framework for Twitter Data using Supervised Classifiers", *International Journal of Computer Applications*, Vol. 5, No. 3, pp. 1-12, 2020.

[20] J. Golbeck, C. Robles, M. Edmondson and K. Turner, "Predicting Personality from Twitter", *Proceedings of IEEE 3rd International Conference on Privacy, Security, Risk and Trust*, pp. 149-156, 2011.

[21] Sepandar D. Kamvar and Jonathan Harris, "We Feel Fine and Searching the Emotional Web", *Proceedings of 4th ACM International Conference on Web Search and Data Mining*, pp. 117-126, 2011.

[22] Johan Bollen, Huina Mao and Xiaojun Zeng, "Twitter Mood Predicts the Stock Market", *Journal of Computational Science*, Vol. 2, No. 1, pp. 1-8, 2011.

[23] Flor Miriam Plaza-Del-Arco, M. Teresa Martín-Valdivia, L. Alfonso Urena-Lopez and Ruslan Mitkov, "Improved Emotion Recognition in Spanish Social Media through Incorporation of Lexical Knowledge", *Future Generation Computer Systems*, Vol. 110, pp. 1000-1008, 2019.

[24] B. Nejat, G. Carenini, and R. Ng, "Exploring Joint Neural Model for Sentence Level Discourse Parsing and Sentiment Analysis", *Proceedings of 18th Annual Meeting on Discourse and Dialogue*, pp. 289-298, 2017.

[25] S.M. Mohammad and P.D. Turney, "Emotions Evoked by Common Words and Phrases: using Mechanical Turk to Create an Emotion Lexicon", *Proceedings of Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pp. 26–34, 2010.