ORIGINAL ARTICLE

Mood detection and prediction using conventional machine learning techniques on COVID19 data

Subhayan Bhattacharya[1] · Abhay Agarwala[1] · Sarbani Roy[1]

Abstract

Emotion detection is a promising field of research in multiple perspectives such as psychology, marketing, network analysis and so on. Multiple models have been suggested over the years for accurate and efficient mood detection. Identifying emotion, or mood, from text has progressed from a simple frequency distribution analysis to far more complicated learning approaches. The main aim of all these text mining and analysis is twofold. First is to categorise existing text into broad classes of emotions, such as happy, sad, angry, surprised and so on. The second aim is to accurately predict the moods of real-time streaming text. The novelty of the work lies in the extensive comparison of nine conventional learning methods with respect to performance metrics precision, recall, F1 and accuracy as well as studying the variance of mood over time using a wide array of moods (25). Using conventional classifiers allow near real-time predictions, can work on considerably less training data, and has the flexibility of feature engineering, as deep learning methods have feature engineering embedded in the model. Since a single line of text can be associated with multiple emotions, this article compares the performance of classifiers in predicting multiple moods for streaming text with likelihood-based ranking. An android application named Citizens' Sense was developed for text collection and analysis. The performance of mood classifiers are tested further using Twitter data related to COVID19. Based on the precision, recall, F1 and accuracy of the classifiers, it can be seen that Random Forest, Decision Tree and Complement Naive Bayes classifiers are marginally better than the other classifiers. The variance of mood over time, and predicted moods for text support this finding.

Keywords Mood detection · Mood prediction · COVID Twitter data · Machine learning

1 Introduction

On an average, 500 million tweets are generated everyday across the world. 2000+ Million users are active on Facebook every month. Instagram records a total of about 9 Million images and posts every day. These numbers provide a glimpse of the variety and velocity at which information is generated on social networks on a day-to-day basis. The online social networks, indeed, are one of the richest sources of data in modern days. Due to the textual nature of a majority of these posts, text mining and analysis has evolved as one of the highlighted research domains with respect to information mining on online social networks.

Text mining and analysis can provide valuable insights about the behaviour of the network (Xu et al. 2013), a group within the network (Yong et al. 2010) or an individual (Bodendorf and Kaiser 2009). The data (text) available on online social networks is of an unstructured nature. For it to be usable, some form of normalisation is required to transform it into structured data. The structured data can then in turn be used for many operations such as classification.

One valuable insight about individuals on online social networks is their mood or mood pattern which can be identified using sentiment analysis (Kharde et al. 2016). Sentiment analysis can be anything between marking an item review on a business-ranking website as positive or negative (Bhatt et al. 2015) to identifying moods from facial expressions of individuals in pictures posted on OSNs (Kulkarni et al. 2009). Mood detection and prediction from textual data is

✉ Sarbani Roy
sarbani.roy@jadavpuruniversity.in

Subhayan Bhattacharya
b.subhayan@yahoo.com

Abhay Agarwala
abhayagarwala011@gmail.com

[1] Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

one such research area which has found wide-spread applications in real-life scenarios.

From predicting emojis based on the real-time text being typed (Ramaswamy et al. 2019) to identifying mental health scenarios from analysing the mood trends of the activity of individuals on OSNs over time (Kang et al. 2016), mood detection and prediction has made the social networking experience better and more efficient. The focus of this paper is to detect mood from archival data and to predict mood for real-time streaming text using classifiers. An android application CitizenSense was developed for data collection and validation where users express opinions/ suggestions/ comments on a wide array of topics. The nature of such posts is predominantly textual and each such post has to be accompanied by at least one (or more) moods/emotions which can be chosen from 25 mood/emotion options. The novelty of this work is the extensive nature of the comparative study of nine conventional learning techniques, and the inclusion of a wide array of moods for classification. The collected data is utilised in comparative analysis of multiple existing conventional learning techniques such as Decision Tree, Support Vector Machine, Multinomial Naive Bayes and so on for mood detection, and to adapt them to predict top three probable moods for real-time streaming data. Two tweet data sets are also considered to study the variance of moods/emotions detected over a period of time.

1.1 Motivation

Online social network is currently a major source of information that can be instrumental in forming and influencing opinions and perspectives. The textual data present on OSNs can be misinterpreted causing unintended confusing and misleading opinions. Associating moods/emotions with textual data can prevent misinterpretation to a substantial degree. Predicting moods can enhance the online social networking experience of the naive user. Predicting the top three moods/emotions theoretically has a higher hit-ratio as well. Classifiers can give a simple and real-time mood detection and prediction.

1.2 Contribution

This work has the following contributions—

- A comparative study of the performance of different text classification algorithms in identifying moods and study the variance in moods over time from Tweet data sets.
- Classification based on a large set of moods (25) that covers an wide array of emotions, identifying three possible moods for each text in real-time streaming data.

- Experimentation on real-life data set for accurate detection of moods for each text.

Figure 1 represents a schematic diagram of the framework utilised in this work for comparative study of mood detection and prediction for textual data.

1.3 Organisation

The rest of the article has been organised as follows—Section 2 provides a literature review of some of the trend-setting and recent research work. Section 3 formally defines the problem that this article is trying to solve, and Section 4 discusses the proposed approach for solving the problem. The experimental setup, experimental results and subsequent analysis are provided in Section 5. Section 6 concludes the paper and highlights future scope of research.

2 Literature survey

There have been multiple recognised work in the field of emotion detection and prediction. There are a variety of approaches such as lexicon classification and neural models. Generating lexicons for emotion detection (Bandhakavi et al. 2017) makes processing text for emotion detection easier. Valence Arousal Dominance Lexicon (Mohammad 2018) is another noteworthy work in lexicon-based mood identification. A supervised learning model for classification of text to mood was suggested in Hasan et al. (2018). A Bayesian classifier for classifying song lyrics is presented in An et al. (2017). In Li et al. (2016) a learning approach towards emotion detection from text is presented. Classification of mood beyond the six classes of Ekman's model has also been explored in Agrawal and An (2012).

A hybrid method with a high prediction accuracy, combining keyword and learning-based model for emotion detection from text was suggested by Binali et al. (2010). Another hybrid model based on lexicon and machine learning for emotion detection from text is given by Dini and Bittar (2016). A regression learning approach to finding mood intensities from text is discussed in Mohammad and Bravo-Marquez (2017). A hybrid model of natural language processing and graph-based learning approach for computing similarity in tweet text with respect to emotions is studied in Summa et al. (2016). A support vector machine and convolutional neural network-based emotion detection framework is proposed by Sen et al. (2017). A combination of biterm topic model and convolutional neural network for emotion detection is discussed in Li et al. (2016).
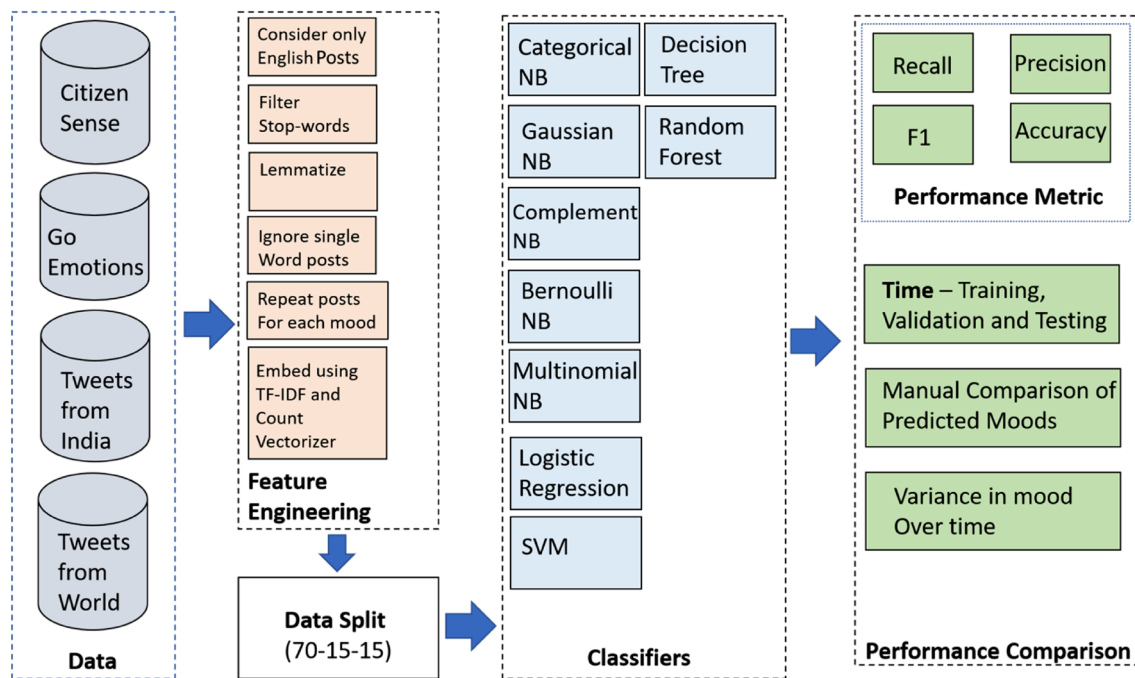
Fig. 1 Framework for comparative study of mood detection and prediction for textual data

Another interesting offshoot of this research field is mood detection and prediction from non-English languages, which poses certain difficulties as not enough corpus or open-source code libraries are available for non-English text. Hussein et al. (2020) proposes a framework for mood detection and prediction for Arabic language. A framework for detection of mood from song lyrics in Bangla has been proposed in Urmi et al. (2020).

There are a number of surveys as well that covers the variety of works done in this domain over the years. Anagnostopoulos et al. (2012) provides an overview of the classifiers and features that have been commonly used in mood detection from text during the years 2000 to 2011. Sailunaz et al. (2018) provides a survey of techniques and technology for mood detection from textual and audio content.

This work is focused on comparing the classifiers with respect to the efficiency and accuracy in predicting moods from text.

3 Problem definition

3.1 Classifying emotion/mood of text corpora

The first problem addressed in this article is a comparative analysis of how some of the existing classification algorithms for textual data performs for mood detection. There are 25 possibilities for moods in the text corpora that has been considered for the experimental results. Given a text corpora $T = \{t_1, t_2, ..., t_n\}$ where t_i is the text associated with a single post, a set of moods $M = \{m_1, m_2, ..., m_p\}$ where m_j is a mood, a set of classifiers $C = \{c_1, c_2, ..., c_q\}$, and a set of similarity metric $S = \{s_1, s2, ..., s_r\}$, find a classifier $c \in C$ such that the similarity $\forall s_k \in S$ is maximised for $c : t_i \rightarrow m_j$.

Each classifier $c_l \in C$ behaves differently, and it is impractical to judge the performance of all classifiers using one similarity metric. So the objective here is to identify a classifier that performs well for all similarity metric $s_k \in S$. As it is not within the scope to test the performance of the classifiers against all possible similarity metrics, thus, a subset S has been chosen for the purpose of this work. For different S, the best classifier identified following the same procedure might vary.

3.2 Predicting emotion/mood of real-time streaming data

The second problem addressed in this article is predicting the emotion/mood associated with textual data.

Given a single unit of textual data τ, a trained classifier c and a set of moods $M = \{m_1, m_2, ..., m_p\}$, find $c : \tau \rightarrow m_i, m_j, m_k, (m_i, m_j, m_k) \in M$ to *maximise likelihood* of (m_i, m_j, m_k).

The aim here is to be able to predict moods for single line of text using a trained classifier such that the predicted moods are as close to the actual moods as possible.

However, as the actual moods might not be available for the single line of text, the problem has been formulated as a maximisation problem. The idea is to maximise the likelihood of the predicted mood.

4 Methodology

4.1 Data definition

– CitizenSense (play 2022)—an android application CitizenSense was developed for data collection. The application was circulated amongst a closed group of people from different streams of work with different demographics. In the application, users can post textual and image contents, accompanied by one or more moods/emotions that can be selected from a set of 25 moods (such as happy, sad, excited, and so on). Data was collected in the application from 10/05/2020 to 28/05/2020. Users were free to post on any topic of their choice. At the end of data collection, a total of approximately 1500 posts were collected, and the cumulative mood count for these posts was approximately 2200.

– Twitter Data for COVID (Lamsal 2020)—data scrapped from Twitter related to COVID19 from across the globe. For the purpose of this research, data for 66 weeks starting from 19th March 2020 till 24th June 2021, with 1000 tweets every 7 days have been considered. These 1000 tweets might contain retweets. This data has not been considered for training, but only for prediction.

– Twitter Data for COVID from India (Lamsal 2020)—data scrapped from Twitter, originating in India, related to COVID19 lockdown. Data for 13 days between 25th March 2020 and 02nd September 2020 with 1000 tweets for each day has been considered. These 1000 tweets might contain retweets. This data has not been considered for training, but only for prediction.

4.2 Feature engineering

The data collected from the android application CitizenSense and the two tweet data sets are noisy data. There are many texts in languages other than English, there are emoticons embedded in the text and so on. Also, the classifiers considered for this work are conventional learning classifiers, and they require a set of features as input for accurate classification. The following steps help identify the features—

- Remove all non-English posts.
- Filter out the stop-words from each of the posts.
- Lemmatise each word in the post to the primitive form (example, running, ran to run)
- Ignore single word posts
- If a post has 'n' moods associated with it, then repeat the post 'n' times, each occurrence associated with one of the 'n' moods.
- Each post is embedded using TF-IDF (term frequency-inverse document frequency) and count vectorizer.

For the purpose of this work, we have utilised multi-class rule-based, Bayesian, and Gaussian classifiers. Binary classifiers are extended for multi-class classifications. The data preparation for all classifiers considered in this work follow the same set of steps as mentioned above. The feature engineering steps have been so designed that they can be generalised and applied to multiple types of classifiers. An alternative approach would be phrase-based representation, but there is no performance improvement (Scott and Matwin 1999). Yet another alternative can be use of hypernyms. However, hypernym representation enhances the performance mainly for rule-based classifiers and have little to no effect on other classifiers (Scott and Matwin 1999).

4.3 Mood detection

From the collected data 70% of data is considered as training data for the classifiers, 15% is considered for validation and the remaining 15% is considered as testing data. For the two tweet data sets, the complete data was used for prediction.

4.3.1 Classifiers

For the purpose of this comparative study, only conventional learning classifiers have been considered. The reason behind selecting conventional learning classifiers is as follows-

- conventional learning algorithms have acceptable performance with a smaller training data set
- conventional learning algorithms require less training time
- conventional learning algorithms require less processing power and resources

Thus, these classifiers have the potential to be utilised to provide near real-time results on devices with limited processing resources and is adaptable for all kinds of data sets. While there exists comprehensive surveys on novel methods

custom designed for detecting emotion from text (Sailunaz et al. 2018) and algorithms for specific emotion detection from text (Aldunate et al. 2018; Hasan et al. 2019) can be found, the aim of this paper is to verify if existing conventional machine learning algorithms can detect and predict moods from a set of large available moods, based on a small training data, with moderate-to-good accuracy. Thus, for the purpose of this comparative study, more problem-specific and custom-designed algorithms have not been considered.

The following classifiers were chosen for classifying the data into moods/emotions

- Naive Bayes classifiers—is a class of supervised learning classifiers that uses conditional probability and independent features to classify text into predefined classes. Naive Bayes classifiers use the bag of word approach, where a document (tweet/post for this work) is considered to be a collection of words, irrespective of the order of words. For this work, five variants of Naive Bayes classifiers were considered, where each variant utilises a different distribution model for classification.

  - Multinomial Naive Bayes—is more suited for classifying text documents based on term frequency. It utilises multinomial distribution.
  - Bernoulli Naive Bayes—is more suited for discrete data where the features are of binary nature. It utilises the multivariate Bernoulli distribution.
  - Categorical Naive Bayes—is more suited for discrete data where the features are categorically distributed.
  - Complement Naive Bayes—is more suited for imbalanced data sets, where the data is not uniformly distributed with respect to classes.
  - Gaussian Naive Bayes—is more suited for continuous data and utilises the Gaussian Distribution for classification.

- Decision tree—uses a set of decision rules for branching to assign data to classes. Each internal node is considered to be a test or decision point, and each leaf node is considered to be a class. Based on the feature values, the tree is traversed from root to leaf, and the class is determined.
- Support vector machine—separates points into classes using a hyperplane where the cumulative distance of the points from the hyperplane is maximised. The number of features for classification determines the dimension of the hyperplane.
- Random Forest—is a collection on uncorrelated, randomly generated Decision Trees which uses bagging and feature randomness. For classification, the result of the random forest is the class predicted by maximum of the Decision Trees. Random Forest classifier determines the features used for classification randomly.
- Logistic Regression—is a predictive analysis classifier that is suitable for dependent variables of binary nature. It can be extended to fit multiple classes as well. The classifier utilises a logistic function for modelling a dependent variable of binary nature.

4.3.2 Classes

The following moods/emotions were available in the CitizenSense application—neutral, anxious, scared, stressed, happy, sad, angry, mad, excited, sorry, proud, curious, lonely, relaxed, nervous, bored, confused, worried, grateful, greedy, hopeless, exhausted, annoyed, calm, motivated.

4.3.3 Performance metric

The testing data is labelled data where each post is associated with one or more moods/emotions. Given a total set of moods M, a post p with actual moods $M_{actual} \subseteq M$, and a set of predicted moods $M_{predicted} \subseteq M$, a true positive is marked $\forall m \in M_{actual} \cap M_{predicted}$, a false negative is marked $\forall m \in M_{actual} \setminus M_{predicted}$, a false positive is marked $\forall m \in M_{predicted} \setminus M_{actual}$, and a true negative is marked $\forall m \in M \setminus (M_{actual} \cup M_{predicted})$.

The following metrics are considered for performance comparison for each of the classifiers -

- $Precision = \frac{TruePositive}{TruePositive + FalsePositive}$
- $Recall = \frac{TruePositive}{TruePositive + FalseNegative}$
- $F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$
- $Accuracy = \sum_{m \in M} \frac{TruePositive}{TotalNumberofPosts}$

Precision, recall and F1 metrics are calculated for each of the 25 moods for each of the nine classifiers that have been used in this comparative analysis. The accuracy is cumulative accuracy over all the moods for each of the nine classifiers.

It should be noted here that for the purpose of this paper, top three of the predicted moods have been considered, and all performance metric calculations are based on this. However, theoretically, considering a higher number of predicted moods could possibly result in better performance. Here, the top three of the predicted moods have been considered as the likelihood of the top 4th mood onward was considerably lower in most of the predictions.

Table 1 Validation—precision (*p*), recall (*r*) and F1 score for moods using logistic regression, multinomial Naive Bayes, Bernoulli Naive Bayes, and Complement Naive Bayes

	Logistic			MultinomialNB			BernoulliNB			ComplementNB		
	r	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1
Angry	0.13	1.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.24	0.31
Excited	0.57	0.66	0.61	0.30	0.87	0.44	0.12	0.83	0.20	0.49	0.38	0.42
Calm	0.95	0.26	0.41	0.98	0.24	0.38	0.98	0.19	0.32	0.69	0.52	0.59
Curious	0.21	0.75	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.23	0.28
Relaxed	0.58	0.41	0.48	0.39	0.38	0.38	0.03	0.33	0.05	0.55	0.29	0.38
Worried	0.22	0.55	0.32	0.03	1.00	0.07	0.03	1.00	0.06	0.39	0.24	0.30
Annoyed	0.45	0.93	0.61	0.04	1.00	0.07	0.03	1.00	0.06	0.42	0.48	0.45
Scared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.15	0.22
Anxious	0.12	1.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.20	0.26
Motivated	0.57	0.31	0.40	0.38	0.61	0.47	0.14	0.71	0.24	0.51	0.35	0.41
Neutral	0.86	0.21	0.34	0.88	0.18	0.31	1.00	0.14	0.24	0.48	0.35	0.40
Sad	0.06	1.00	0.11	0.00	0.00	0.00	0.05	1.00	0.10	0.44	0.32	0.37
Greedy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Stressed	0.14	0.30	0.19	0.11	1.00	0.20	0.04	1.00	0.07	0.43	0.31	0.36
Bored	0.17	0.75	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.29	0.39
Happy	0.94	0.30	0.46	0.99	0.26	0.41	0.97	0.25	0.40	0.65	0.52	0.58
Confused	0.11	1.00	0.20	0.07	1.00	0.13	0.00	0.00	0.00	0.26	0.44	0.33
Exhausted	0.50	0.60	0.55	0.47	0.84	0.60	0.00	0.00	0.00	0.66	0.40	0.49
Hopeless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.44	0.47
Sorry	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.40	0.40
Mad	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.60	0.55
Nervous	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.40	0.44
Grateful	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71	0.34	0.47
Proud	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.50	0.46
Lonely	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.25	0.35

4.4 Prediction

Based on the performance analysis using the metrics provided in the above section, the best performing classifier is chosen and that classifier is used to predict the top three possible moods for a post ordered by maximum likelihood.

5 Experimental results and discussion

Tables 1 and 2 represent the precision, recall and F1 scores of the nine classifiers that were used for mood detection using the validation data. All the classifiers received the same pre-processed input. Similarly, Tables 3 and 4 represent the precision, recall and F1 scores of the nine classifiers using the testing data, which comprises of 15% of the total data set.

In the data collected using the CitizenSense application, not all moods have the same frequency. On the contrary, some of the moods appear a lot more compared to the others. Top five common moods or emotions in the data set are neutral, happy, calm, excited, and motivated, in decreasing order of frequency. The bottom five common moods are nervous, greedy, proud, sorry, and lonely, in decreasing order of frequency. Table 2 shows the exact frequency of moods in the data set. Due to the lack of enough training data for the moods that are seldom associated with the posts, the probability of mis-classification also increases.

As can be seen from the results, the precision, recall and F1 scores are considerably low for all the classifiers, in both validation and testing. Also, the scores vary greatly depending on the frequency at which the moods appear in the collected data set.

Based on Tables 1 and 2, Complement Naive Bayes classifier, Logistic Regression classifier, Gaussian Naive Bayes classifier, Decision Tree classifier, and Random Forest classifier provides better results than Multinomial Naive Bayes classifier, Bernoulli Naive Bayes classifier, Categorical Naive Bayes classifier, and Support Vector Machine. These tables represent the results from the validation part of the data set. Tables 3 and 4 give similar results for the testing part of the data set. Tallying the data

Table 2 Validation—precision (*p*), recall (*r*) and F1 score for moods using categorical Naive Bayes, Gaussian Naive Bayes, random forest, support vector machine (SVM), and Decision Tree

	CategoricalNB			GaussianNB			RandomForest-Classifier			SVM			DecisionTree		
	r	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1
Angry	0	0	0	0.82	0.05	0.1	0.29	0.8	0.42	0	0	0	0.79	0.04	0.08
Excited	0	0	0	0.49	0.61	0.54	0.66	0.51	0.57	0.61	0.43	0.5	0.43	0.59	0.49
Calm	1	0.23	0.38	0.36	0.56	0.44	0.77	0.41	0.54	0.94	0.26	0.41	0.42	0.39	0.4
Curious	0	0	0	0.56	0.63	0.59	0.44	0.37	0.4	0	0	0	0.26	0.67	0.38
Relaxed	0	0	0	0.38	0.34	0.36	0.73	0.48	0.58	0.35	0.63	0.45	0.44	0.53	0.48
Worried	0	0	0	0.31	0.53	0.39	0.36	0.5	0.42	0.32	0.25	0.28	0.35	0.8	0.48
Annoyed	0	0	0	0.78	0.08	0.15	0.31	0.47	0.38	0.31	0.53	0.39	0.88	0.11	0.2
Scared	0	0	0	0.38	0.33	0.35	0.42	0.71	0.53	0	0	0	0.22	1	0.36
Anxious	0	0	0	0.37	0.35	0.36	0.4	0.73	0.52	0	0	0	0.43	0.55	0.48
Motivated	0	0	0	0.42	0.62	0.5	0.45	0.27	0.33	0.43	0.71	0.54	0.32	0.62	0.42
Neutral	1	0.16	0.28	0.21	0.52	0.3	0.82	0.29	0.43	0.92	0.21	0.34	0.51	0.39	0.44
Sad	0	0	0	0.32	0.21	0.26	0.3	0.43	0.35	0	0	0	0.38	0.5	0.43
Greedy	0	0	0	0.67	0.32	0.43	0	0	0	0	0	0	0.2	0.67	0.31
Stressed	0	0	0	0.18	0.3	0.22	0.53	0.4	0.46	0.16	0.5	0.24	0.35	0.6	0.44
Bored	0	0	0	0.38	0.3	0.33	0.41	0.39	0.4	0.47	0.78	0.58	0.47	0.3	0.37
Happy	1	0.29	0.45	0.16	0.75	0.27	0.8	0.3	0.44	0.94	0.28	0.44	0.48	0.68	0.56
Confused	0	0	0	0.21	0.5	0.29	0.41	0.6	0.49	0	0	0	0.29	0.5	0.36
Exhausted	0	0	0	0.39	0.41	0.4	0.65	0.3	0.41	0.77	0.49	0.6	0.56	0.7	0.62
Hopeless	0	0	0	0.43	0.46	0.44	0.35	0.67	0.46	0	0	0	0.43	0.69	0.53
Sorry	0	0	0	0.29	0.33	0.31	0.67	0.67	0.67	0	0	0	0.75	0.6	0.67
Mad	0	0	0	0.31	0.27	0.29	0.38	1	0.55	0.08	1	0.15	0.45	0.83	0.59
Nervous	0	0	0	0.56	0.36	0.43	0.38	0.5	0.43	0	0	0	0.33	1	0.5
Grateful	0	0	0	0.18	0.27	0.21	0.3	0.33	0.32	0	0	0	0.55	0.92	0.69
Proud	0	0	0	0.5	0.8	0.62	0.29	1	0.44	0	0	0	0.4	1	0.57
Lonely	0	0	0	0.67	0.31	0.42	0.25	1	0.4	0	0	0	0	0	0

from these two tables, it can be safely commented that there is no overfitting or underfitting.

Figure 2 represents the total number the posts corresponding to each of the moods. Figure 3 represents the cumulative accuracy of the classifiers over the test data set. Since the split of training and testing data is random, the average of five iterations is considered for calculating accuracy. From this figure also, it can be seen that Random Forest, Decision Tree and Complement Naive Bayes classifiers have the highest accuracy, followed by Multinomial Naive Bayes.

Taking into account all the four performance metric under consideration, Random Forest classifier, Decision Tree, and Complement Naive Bayes have the best average performance as they have higher number of moods with nonzero precision, recall and F1 score and accuracy. For a nonzero score, at least some true positives are present, which means that at least one post was correctly classified for each of these moods, even if the moods are repeated very less in the data set. The performance of these three classifiers are further

studied using the two real-life tweet data sets. The training, validation and testing time for each of the classifiers have been presented in Table 5.

Based on the performance of the different classifiers, Decision Tree, Complement Naive Bayes, and the Random Forest classifier were chosen for prediction. For prediction, 100 posts were chosen randomly from an open source data set[1]. Some of the positive and negative results are being listed here. The original data sets have 1 mood associated with them out of a possible 28 moods. The list of 28 moods for the given dataset is similar to the 25 moods considered for the purpose of this work. Thus, this data set has been selected. However, as the aim here is to identify the top three moods in decreasing order of likelihood, 100 volunteers were requested to assign three moods in decreasing order of likelihood to each of the sentences. Then, the mood with maximum frequency over 100 volunteers responses for sentence 1 was chosen as mood with maximum likelihood

---

[1] https://github.com/google-research/google-research/tree/master/goemotions/data

Table 3 Prediction—Precision (*p*), Recall (*r*) and F1 score for moods using logistic regression, multinomial Naive Bayes, Bernoulli Naive Bayes, and Complement Naive Bayes

	Logistic			MultinomialNB			BernoulliNB			ComplementNB		
	r	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1
Angry	0.12	1.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.29	0.38
Excited	0.54	0.51	0.52	0.21	0.82	0.34	0.13	0.50	0.20	0.45	0.35	0.40
Calm	0.92	0.29	0.44	0.97	0.24	0.38	0.99	0.24	0.38	0.74	0.46	0.57
Curious	0.13	0.67	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.41	0.38
Relaxed	0.73	0.41	0.52	0.42	0.41	0.41	0.10	0.80	0.18	0.68	0.38	0.49
Worried	0.29	0.40	0.33	0.14	1.00	0.25	0.07	1.00	0.14	0.48	0.33	0.39
Annoyed	0.27	0.64	0.38	0.00	0.00	0.00	0.05	1.00	0.09	0.50	0.32	0.39
Scared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46	0.32	0.37
Anxious	0.06	1.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.20	0.23
Motivated	0.49	0.31	0.38	0.29	0.55	0.38	0.08	0.60	0.13	0.50	0.44	0.47
Neutral	0.77	0.16	0.26	0.89	0.19	0.31	1.00	0.16	0.27	0.66	0.35	0.46
Sad	0.12	1.00	0.21	0.08	1.00	0.15	0.00	0.00	0.00	0.42	0.33	0.37
Greedy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.08	0.09
Stressed	0.17	0.33	0.22	0.05	1.00	0.10	0.00	0.00	0.00	0.38	0.33	0.36
Bored	0.33	0.56	0.42	0.05	1.00	0.10	0.00	0.00	0.00	0.68	0.38	0.49
Happy	0.92	0.34	0.49	1.00	0.25	0.40	0.99	0.27	0.42	0.66	0.51	0.58
Confused	0.13	1.00	0.23	0.00	0.00	0.00	0.05	1.00	0.09	0.33	0.33	0.33
Exhausted	0.52	0.55	0.53	0.48	0.63	0.55	0.04	0.33	0.06	0.54	0.30	0.39
Hopeless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.64	0.35	0.45
Sorry	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.30	0.37
Mad	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.57	0.47	0.52
Nervous	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.31	0.40
Grateful	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.64	0.27	0.38
Proud	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.30	0.32
Lonely	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.24	0.36

for sentence 1, mood with second maximum frequency over 100 volunteers responses for sentence 1 was chosen as mood with second maximum likelihood for sentence 1 and so on.

Table 6 shows the predictions using Random Forest classifier alongside the actual moods. As can be seen, the classifier is able to capture the moods for some of the posts, while completely misses the mood for some of the other posts. Another point to note is that most of the posts have "Neutral" as one of the top three predicted moods, for Random Forest Classification.

Similarly, Tables 7 and 8 represent the predictions made by Complement Naive Bayes classifier and Decision Tree classifier alongside the actual moods. The same posts have been considered in all three tables to maintain generalisability. It is interesting to note that, in contrast to Random Forest classifier, both these classifiers do not show a bias towards the mood "Neutral".

As can be seen from these tables, while some of the predictions are able to capture the essence of the sentence, others are not accurate. The poor prediction of the classifiers can be partially attributed to the high frequency of some moods in the training data and low frequency of some of the other moods. Also, the classifiers are unable to capture some of the intricacies and figures of speech present in the sentences, like sarcasm and so on. Thus, the prediction supports the observations made on the test data.

As an alternative approach, we considered using Stratified Sampling for a more uniform distribution of moods in the training data. The results observed during testing using stratified sampling are recorded in Appendix D. It can be seen that the cumulative accuracy of all the classifiers (shown in Fig. 21) is considerably lower than the cumulative accuracy as shown in Fig. 3. However, the recall, precision and F1 score are quite high for all the moods for all the classifiers (as shown in Tables 12 and 13). Another factor contributing towards the anomalies in the prediction is non-textual elements in the posts, such as emojis. For example, the post "I'm so sorry 😔" is predicted to have the moods

Table 4 Prediction—precision (*p*), recall (*r*) and F1 score for moods using Categorical Naive Bayes, Gaussian Naive Bayes, random forest, support vector machine(SVM), and Decision Tree

	CategoricalNB			GaussianNB			RandomForest-Classifier			SVM			DecisionTree		
	r	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1
Angry	0.00	0.00	0.00	0.92	0.05	0.09	0.42	0.42	0.42	0.00	0.00	0.00	0.87	0.05	0.09
Excited	0.00	0.00	0.00	0.53	0.58	0.55	0.52	0.38	0.44	0.59	0.51	0.55	0.48	0.68	0.56
Calm	1.00	0.21	0.35	0.38	0.56	0.45	0.68	0.32	0.44	0.99	0.28	0.44	0.57	0.58	0.58
Curious	0.00	0.00	0.00	0.17	0.25	0.20	0.47	0.32	0.38	0.00	0.00	0.00	0.33	0.67	0.44
Relaxed	0.00	0.00	0.00	0.43	0.55	0.48	0.47	0.33	0.39	0.43	0.53	0.48	0.24	0.53	0.33
Worried	0.00	0.00	0.00	0.16	0.60	0.25	0.40	0.50	0.44	0.52	0.52	0.52	0.54	0.67	0.60
Annoyed	0.00	0.00	0.00	0.74	0.10	0.17	0.42	0.46	0.44	0.14	0.25	0.18	0.81	0.10	0.19
Scared	0.00	0.00	0.00	0.22	0.18	0.20	0.71	0.75	0.73	0.00	0.00	0.00	0.38	0.75	0.50
Anxious	0.00	0.00	0.00	0.27	0.29	0.28	0.56	0.63	0.59	0.00	0.00	0.00	0.31	0.42	0.36
Motivated	0.00	0.00	0.00	0.33	0.76	0.46	0.53	0.30	0.38	0.36	0.71	0.48	0.35	0.68	0.46
Neutral	1.00	0.15	0.26	0.21	0.33	0.26	0.70	0.26	0.38	0.90	0.21	0.34	0.45	0.39	0.42
Sad	0.00	0.00	0.00	0.22	0.25	0.24	0.39	0.54	0.45	0.00	0.00	0.00	0.38	0.67	0.48
Greedy	0.00	0.00	0.00	0.83	0.50	0.63	0.36	0.57	0.44	0.00	0.00	0.00	0.33	0.60	0.43
Stressed	0.00	0.00	0.00	0.40	0.47	0.43	0.36	0.40	0.38	0.25	0.83	0.38	0.43	0.82	0.56
Bored	0.00	0.00	0.00	0.37	0.26	0.30	0.35	0.54	0.42	0.14	0.75	0.23	0.37	0.44	0.40
Happy	1.00	0.20	0.33	0.14	0.61	0.22	0.87	0.29	0.43	0.93	0.26	0.41	0.52	0.60	0.56
Confused	0.00	0.00	0.00	0.20	0.50	0.29	0.35	0.37	0.36	0.00	0.00	0.00	0.41	0.64	0.50
Exhausted	0.00	0.00	0.00	0.33	0.45	0.38	0.67	0.26	0.38	0.86	0.51	0.64	0.47	0.62	0.53
Hopeless	0.00	0.00	0.00	0.53	0.60	0.56	0.31	0.33	0.32	0.00	0.00	0.00	0.29	1.00	0.44
Sorry	0.00	0.00	0.00	0.25	1.00	0.40	0.67	1.00	0.80	0.00	0.00	0.00	0.40	0.50	0.44
Mad	0.00	0.00	0.00	0.80	0.42	0.55	0.33	0.80	0.47	0.00	0.00	0.00	0.45	0.71	0.56
Nervous	0.00	0.00	0.00	0.33	0.67	0.44	0.10	0.25	0.14	0.00	0.00	0.00	0.29	1.00	0.44
Grateful	0.00	0.00	0.00	0.38	0.33	0.36	0.21	0.75	0.33	0.00	0.00	0.00	0.33	0.56	0.42
Proud	0.00	0.00	0.00	0.30	1.00	0.46	0.22	1.00	0.36	0.00	0.00	0.00	0.13	1.00	0.22
Lonely	0.00	0.00	0.00	0.50	0.43	0.46	0.40	1.00	0.57	0.00	0.00	0.00	0.60	0.75	0.67
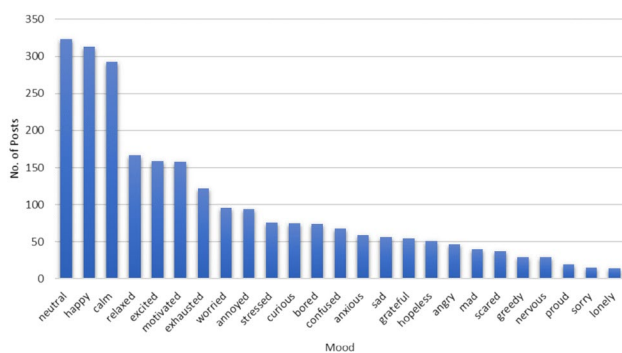
Fig. 2 No. of posts versus mood

Fig. 3 Cumulative accuracy of classifiers

Sorry, Neutral, and Mad in decreasing order of likelihood. This prediction is accurate. However, when the emoji is considered, which was a part of the post before pre-processing, the emotion of the post changes which cannot be captured by the classifiers.
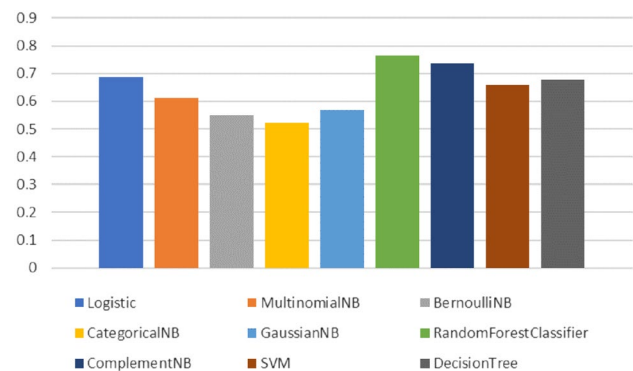
Table 5 represents the execution time for each of the classifiers. This execution time is in seconds for the Citizens' Sense data set where 70% of the data is considered for training, 15% for validation and 15% for testing. The tests were executed on a 2.7GHz processor with 12 GB of memory.

Table 5 Time taken (in seconds) by each Classifier for Training and Testing

Classifier	Training time	Validation time	Testing time
Logistic	1.707196	0.355261	0.276393
Multinomial Naive Bayes	0.041612	0.278217	0.288085
Bernoulli Naive Bayes	0.061874	0.542308	0.523950
Categorical Naive Bayes	1.568812	2.567296	2.418318
Gaussian Naive Bayes	0.090244	0.594380	0.676466
Random Forest	5.873815	3.849990	3.206187
Complement Naive Bayes	0.040305	0.276880	0.272030
Support Vector Machine	24.113053	1.112772	1.108734
Decision Tree	2.055536	0.207499	0.214717

From this table, it can be seen that the Naive Bayes classifiers in general have a lower training time, followed by Logistic, Decision Tree, Random Forest and support vector machine in increasing order of training time. For testing time as well, the Naive Bayes classifiers have lower testing times with the exception of Categorical Naive Bayes. Thus, combining both the accuracy and the execution times, Complement Naive Bayes seems to be the best amongst the classifiers for prediction of mood/emotion from text.

5.1 Mood identification with Twitter data

The two Twitter data sets described in Section IV A were used for identifying moods of tweets related to COVID19. COVID19 has been a life-changing event that has affected countries from all over the world, and netizens took to Twitter to express their views and opinions for the same. Multiple events in the timeline effects the views and opinions expressed on Twitter. Events like the onset of first wave of infections, worldwide lockdown, adverse effect on economy, research on vaccines, invention of vaccines, phases of vaccination, second wave of infections and so on. There must be highs and lows in the tweets from both all over the world as well as the tweets originating from India. The idea is to identify these moods and see if they are apt with the actual timeline.

Figure 4 shows some of the major events in the COVID19 timeline in India. This timeline is not exhaustive, but gives an idea of the inter-weaved nature of positive and negative events for the past year. The COVID19 timeline with respect to the world is shown in Fig. 5. This timeline is also not exhaustive in nature but highlights some of the key developments.

There are some key contrasts in these two timelines. The European countries and the United States of America registered their first cases of COVID a lot earlier than

Table 6 Random Forest Classification - Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	Predicted moods	Actual moods
I just need a quiet place to hide	Exhausted, Calm, Neutral	Exhausted, Stressed, Sad
I love that quote, thank you	Calm, Neutral, Happy	Grateful, Happy, Relaxed
I feel you mate	Mad, Annoyed, Neutral	Neutral, Happy, Relaxed
Because of this bad situation some people are afraid of becoming youtubers because of this	Scared, Neutral, Annoyed	Worried, Annoyed, Anxious
How dare you! Don't you know seatbelts cause autism!	Worried, Sad, Scared	Angry, Mad, Annoyed
Help, help, I'm being repressed!	Exhausted, Annoyed, Neutral	Stressed, Anxious, Nervous
Wow, that's scary, she's so young and healthy. I hate that.	Calm, Happy, Neutral	Scared, Worried, Sad

Table 7 Complement Naive Bayes Classification - Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	Predicted moods	Actual moods
I just need a quiet place to hide	Angry, Hopeless, Sad	Exhausted, Stressed, Sad
I love that quote, thank you	Happy, Motivated, Relaxed	Grateful, Happy, Relaxed
I feel you mate	Exhausted, Calm, Bored	Neutral, Happy, Relaxed
Because of this bad situation some people are afraid of becoming youtubers because of this	Worried, Nervous, Scared	Worried, Annoyed, Anxious
How dare you! Don't you know seatbelts cause autism!	Annoyed, Anxious, Motivated	Angry, Mad, Annoyed
Help, help, I'm being repressed!	Exhausted, Happy, Neutral	Stressed, Anxious, Nervous
Wow, that's scary, she's so young and healthy. I hate that	Mad, Calm, Anxious	Scared, Worried, Sad

Table 8 Decision Tree Classification - Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	Predicted moods	Actual moods
I just need a quiet place to hide	Angry, Exhausted, Annoyed	Exhausted, Stressed, Sad
I love that quote, thank you	Angry, Annoyed, Neutral	Grateful, Happy, Relaxed
I feel you mate	Angry, Annoyed, Neutral	Neutral, Happy, Relaxed
Because of this bad situation some people are afraid of becoming youtubers because of this	Angry, Sorry, Sad	Worried, Annoyed, Anxious
How dare you! Don't you know seatbelts cause autism!	Angry, Annoyed, Neutral	Angry, Mad, Annoyed
Help, help, I'm being repressed!	Angry, Annoyed, Confused	Stressed, Anxious, Nervous
Wow, that's scary, she's so young and healthy. I hate that	Angry, Annoyed, Neutral	Scared, Worried, Sad
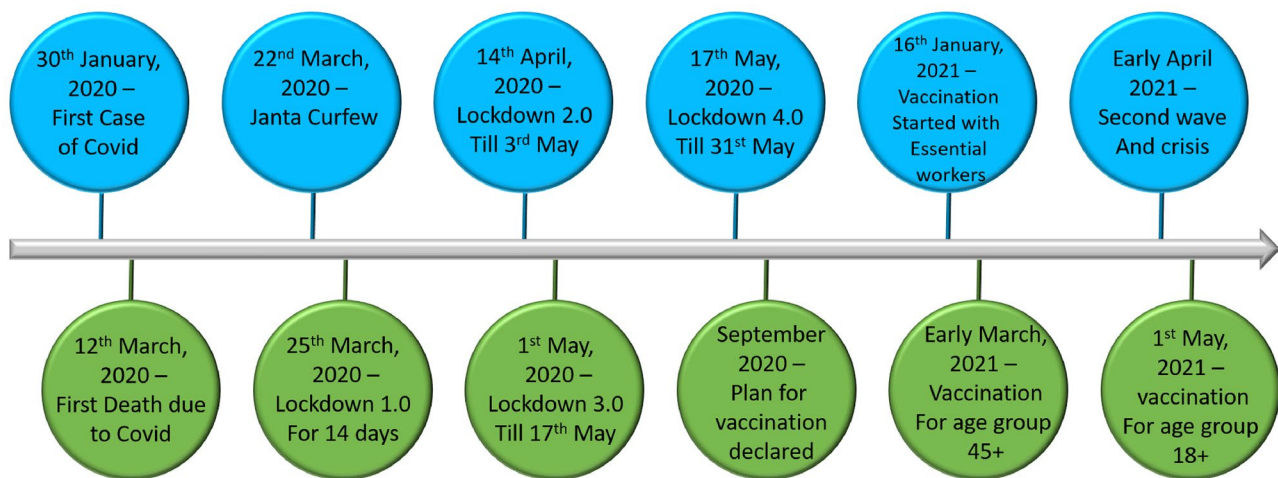
Fig. 4 Brief timeline of COVID19 in India
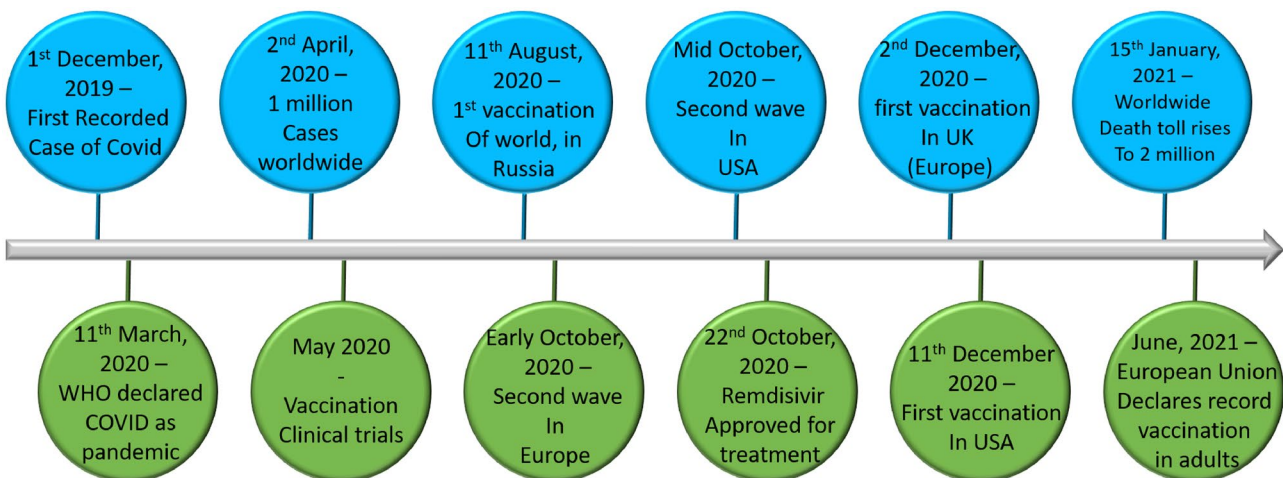
Fig. 5 Brief timeline of COVID19 in the world

India. The first wave of COVID patients were also earlier in these places compared to India. However, lockdown was imposed at a similar time all over the world. The first wave of high number of patients with steep peaks in India was almost coincidental with the second wave of high number of patients in European countries. Vaccine production and

vaccination were also started earlier in Russia, the UK, the USA and other European countries compared to India. These chain of events can have a serious impact on the moods depicted in the tweets used for this research.

Based on the results shown in Tables 3 and 4, the Complement Naive Bayes, Decision Tree, and Random Forest classifiers were considered to give the best results. Thus, these three classifiers were considered for detecting the moods in the tweet data sets as well. All the pre-processing and processing steps used are the same as before. The top three predictions for each mood is considered and a cumulative

count is considered for the total of 1000 tweets for a day. This was repeated for the complete data set (13 days for India data set and 66 days for the World data set). The data is then plotted with the days on *x*-axis and frequency on *y*-axis.

5.1.1 Decision tree

Figure 6 shows the mood classification of tweets from India. Figure 7 shows the mood classification of tweets from all over the world using Decision Trees. For tweets

(a) Non-Stratified (b) Stratified

Fig. 6 Tweets from India classified using Decision Tree classifier

(a) Non-Stratified (b) Stratified
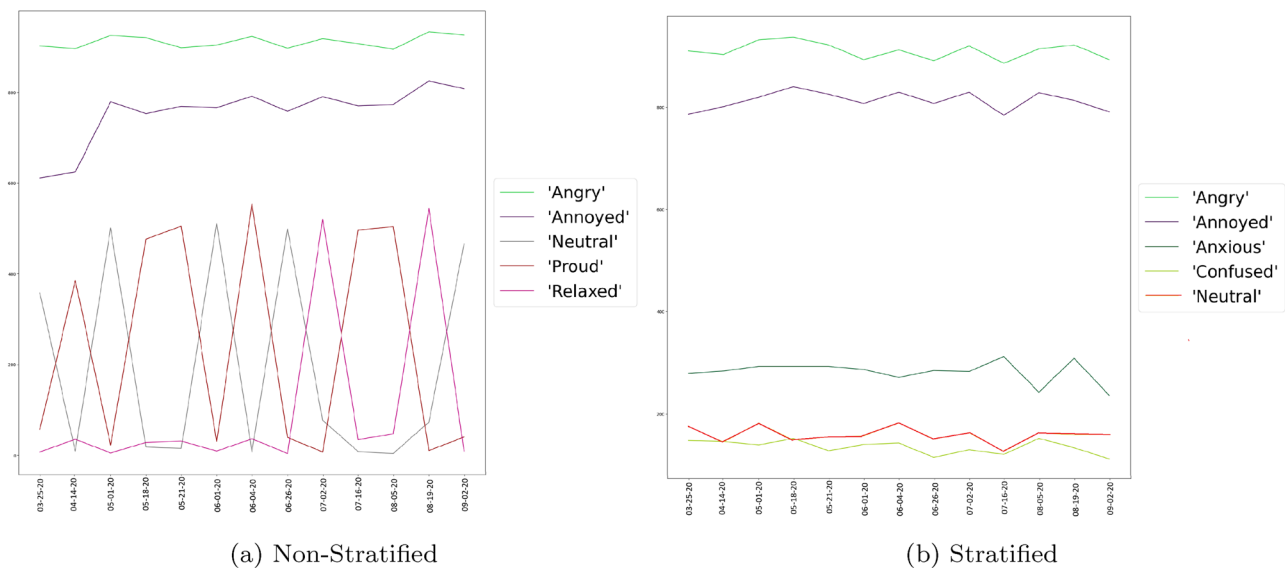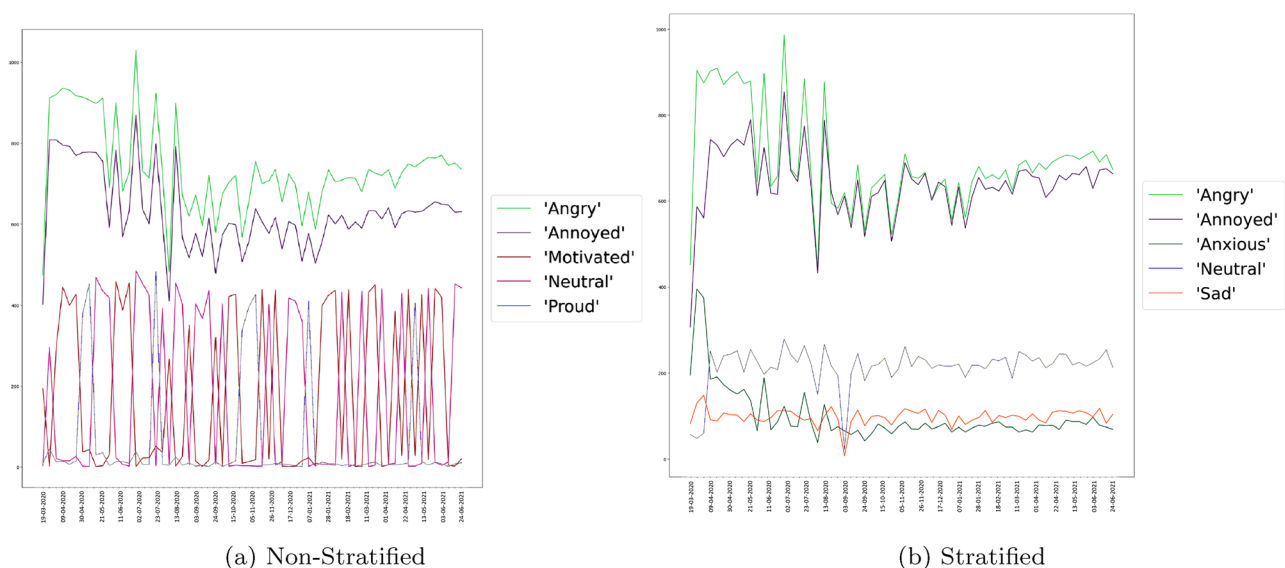
Fig. 7 Tweets from all over the world classified using Decision Tree classifier, data points are 7 days apart

from India, the top 5 moods in terms of maximum likelihood are Angry, Annoyed, Neutral, Proud, and Relaxed for non-stratified sampling and Angry, Annoyed, Anxious, Confused, and Neutral for stratified sampling. For tweets originating from all over the world, the top 5 moods are Angry, Annoyed, Motivated, Neutral, and Proud for non-stratified sampling and Angry, Annoyed, Anxious, Neutral, and Sad for stratified sampling. Throughout both these timelines, Angry and Annoyed are the top two moods. Although this is kind of representative of the overall mood during timeline, it is interesting to note that even in Table 11, Decision Tree predicts Angry and Annoyed frequently. Decision Tree performs best with Binary decisions, and since Angry and Annoyed are the first two moods alphabetically, and thus, Decision Tree predicted these two most frequently. This is a shortcoming of the Decision Tree Classifier. Thus, although Decision Tree is apparently immune to the bias of some mood represented a lot more frequently than others in the training data, in reality, it only predicts two out of the 25 possible moods frequently irrespective of the actual text. Another point to note is that for stratified sampling, there are considerably less spikes, which might indicate that stratified sampling is not optimal for identifying sharp changes in mood.

5.1.2 Random forest

Figure 8 represents the cumulative count of each mood plotted against the dates for tweets from India using non-stratified and stratified sampling. The top 5 moods in terms of maximum likelihood throughout the timeline are Calm, Happy, Motivated, Neutral, and Worried for non-stratified

sampling and Curious, Motivated, Neutral, Scared, and Worried for stratified sampling. Similarly, Fig. 9 represents the same for tweets from all over the world. The top 5 moods in terms of maximum likelihood are Anxious, Happy, Motivated, Neutral, and Worried for non-stratified sampling and Anxious, Curious, Neutral, Proud, and Scared for stratified sampling. Similar to classification using Decision Tree, there are considerably less spikes in the plot. There are a lot of spikes in the plot for tweets from all over the world. However, most of these spikes cannot be correlated with the timeline as shown in Fig. 5 and even the extended timeline considering all the events not listed in this representative timeline. Another point to note, as the number of trees formed and number of variables taken for forming the tree is random in Random Forest classification, each new iteration for predicting the moods associated with the same mood might yield different results, and thus, a whole new plot. Some of these plots/predictions might be more accurate than others, but there is no consistently good behaviour.

5.1.3 Complement Naive Bayes

Figure 10 represents the cumulative count of each mood plotted against the dates for tweets from India and Fig. 11 the same for tweets from all over the world, as predicted by Complement Naive Bayes classifier. The top moods over the timeline for tweets from India are Confused, Curious, Neutral, Scared, and Worried for non-stratified sampling and Annoyed, Grateful, Neutral, Proud, and Worried for stratified sampling. Similarly for tweets from all over the world, the top moods are Confused, mad, Neutral, Proud,
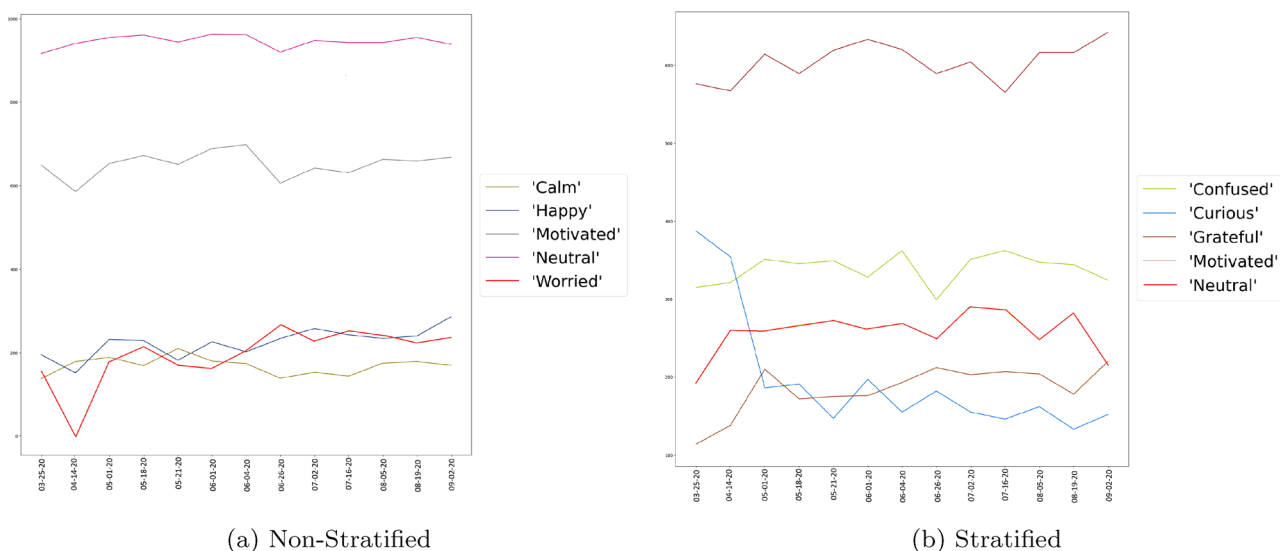

(a) Non-Stratified   (b) Stratified

Fig. 8 Tweets from India classified using random forest classifier

(a) Non-Stratified

(b) Stratified

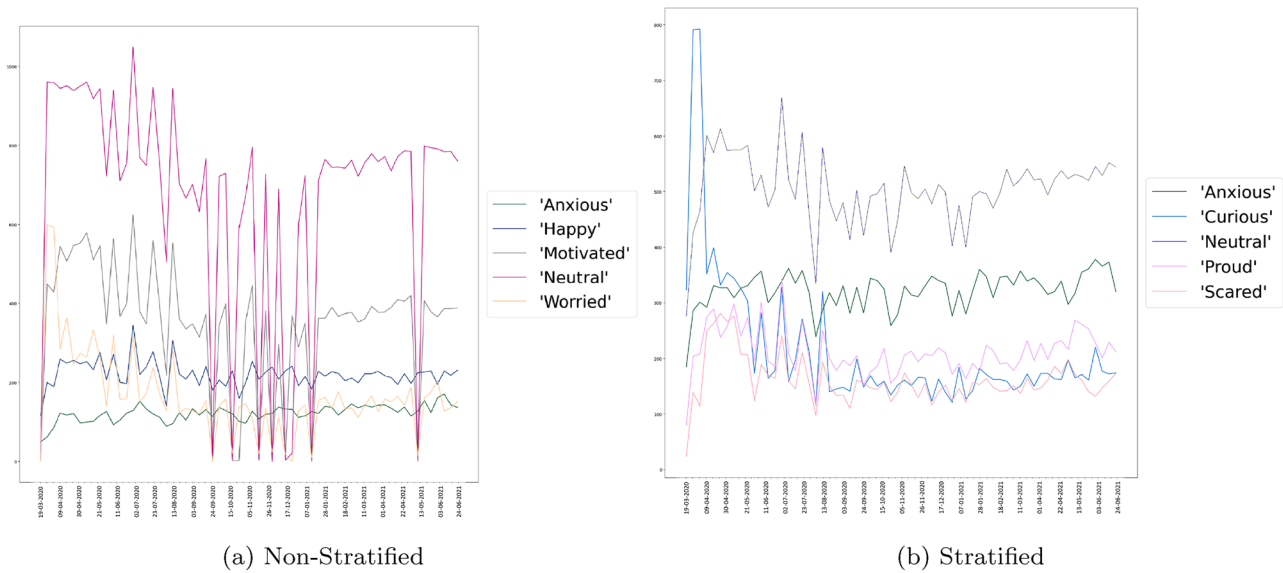**Fig. 9** Tweets from all over the world classified using random forest classifier, data points are 7 days apart

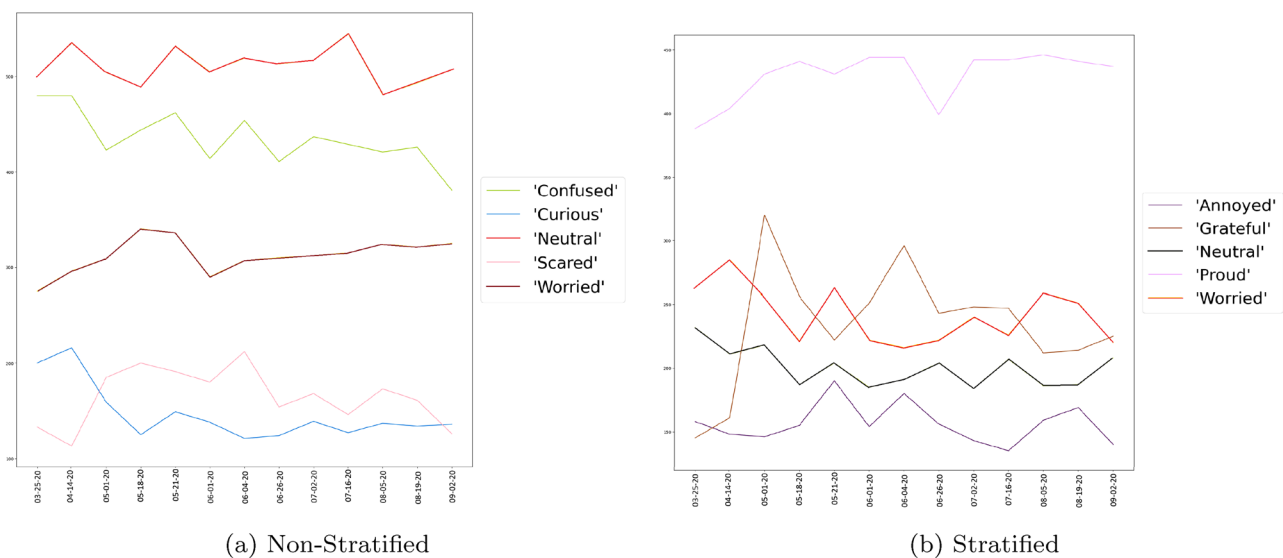(a) Non-Stratified

(b) Stratified

Fig. 10 Tweets from India classified using Complement Naive Bayes classifier

and Worried for non-stratified sampling and Angry, Happy, Mad, Proud, and Scared for stratified sampling. For this classifier, there are spikes and drops as well. Also, although the cumulative count of each mood is consistent throughout the timeline, there are no outright outliers, as is the case of Decision Tree. Rather, all the moods are moderately represented in terms of cumulative count. Thus, Complement Naive Bayes has more consistent predictions and the top moods more or less capture the actual real-life mood in the period of time covered by the data set.

It needs to be noted here that 1000 tweets per day were considered for these experiments. Considering a different number of tweets can result in a different plot using the same classifiers. The idea of the experiment is to study the mood classification using classifiers and analyse the efficiency of the results using different metrics and by correlating to real-life events. From all the experiments conducted, it is safe to assume that classifiers provide inefficient classification for moods from text. However, the classifications are moderately indicative or representative of the actual mood, although they might not be accurate.
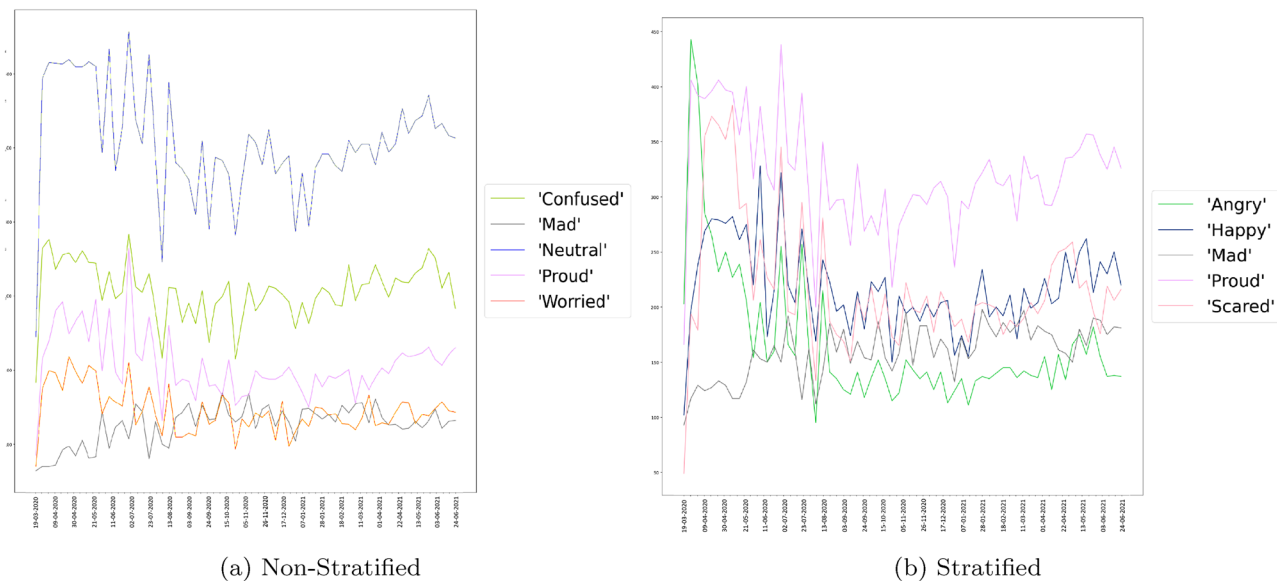
(a) Non-Stratified

(b) Stratified

Fig. 11 Tweets from all over the world classified using Complement Naive Bayes classifier, Data points are 7 days apart

6 Conclusion and future scope

From the comparative analysis of nine basic classifiers, that lack feedback-based learning, it can be observed from the experimental results that the performance of the classifiers is average. None of the classifiers can be termed as distinctively superior to the other classifiers with respect to all four of the performance metrics considered. However, the Complement Naive Bayes classifier can be deemed to be the best of the lot with the highest average performance. Complement Naive Bayes classifier can also be used for predicting moods for real-time streaming data with moderate accuracy, but it is unable to process intricacies like collocation, sarcasm, irony and so on. The execution time for the classifiers can guarantee a near real-time experience with respect to mood prediction. Based on the analysis of the tweet data over an extended time period, Complement Naive Bayes identifies mood variations better than Random Forest and Decision Tree, but there are some inconsistencies in the identification as well.

As a future scope, the classifiers can be enhanced to identify phrasing, sarcasm, figures of speech and so on for better classification. These simple classifiers can then be used for more complicated tasks such as studying mood patterns in individuals on social network for clinical diagnosis of depression, anxiety, distress and so on. Also, the performance of deep learning algorithms on classifying text based on moods, including the intricacies of the language, can also be studied.

Appendix: Predicted moods

This section contains Posts, their predicted moods considering the mood Neutral and their predicted mood not considering the mood neutral using Random Forest, Decision Tree and Complement Naive Bayes classifiers. These results have been included in appendix for the sake of completeness of the experimental results presented (Tables 9, 10, 11).

Table 9 Random Forest Classification—Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	With neutral	Without neutral
I just need a quiet place to hide	Exhausted, Calm, Neutral	Exhausted, Calm, Stressed
I love that quote, thank you	Calm, Neutral, Happy	Relaxed, Calm, Happy
I feel you mate	Mad, Annoyed, Neutral	Annoyed, Mad, Exhausted
Because of this bad situation some people are afraid of becoming YouTubers because of this	Scared, Neutral, Annoyed	Worried, Scared, Sad
How dare you! Don't you know seatbelts cause autism!	Worried, Sad, Scared	Motivated, Worried, Happy
Help, help, I'm being repressed!	Exhausted, Annoyed, Neutral	Exhausted, Calm, Happy
Wow, that's scary, she's so young and healthy. I hate that	Calm, Happy, Neutral	Motivated, Calm, Happy

Table 10 Complement Naive Bayes Classification—Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	With neutral	Without neutral
I just need a quiet place to hide	Angry, Hopeless, Sad	Hopeless, Sad, Angry
I love that quote, thank you	Happy, Motivated, Relaxed	Motivated, Relaxed, Happy
I feel you mate	Exhausted, Calm, Bored	Exhausted, Bored, Calm
Because of this bad situation some people are afraid of becoming YouTubers because of this	Worried, Nervous, Scared	Nervous, Worried, Scared
How dare you! Don't you know seatbelts cause autism!	Annoyed, Anxious, Motivated	Anxious, Excited, Motivated
Help, help, I'm being repressed!	Exhausted, Happy, Neutral	Exhausted, Proud, Happy
Wow, that's scary, she's so young and healthy. I hate that	Mad, Calm, Anxious	Mad, Anxious, Calm

Table 11 Decision Tree Classification—Posts and the predicted moods. Mood 1, Mood 2, and Mood 3 are in decreasing order of likelihood

Post	With neutral	Without neutral
I just need a quiet place to hide	Angry, Exhausted, Annoyed	Annoyed, Exhausted, Angry
I love that quote, thank you	Angry, Annoyed, Neutral	Annoyed, Angry, Calm
I feel you mate	Angry, Annoyed, Neutral	Annoyed, Anxious, Angry
Because of this bad situation some people are afraid of becoming YouTubers because of this	Angry, Sorry, Sad	Worried, Scared, Sad
How dare you! Don't you know seatbelts cause autism!	Angry, Annoyed, Neutral	Worried, Confused, Angry
Help, help, I'm being repressed!	Angry, Annoyed, Confused	Annoyed, Exhausted, Angry
Wow, that's scary, she's so young and healthy. I hate that	Angry, Annoyed, Neutral	Annoyed, Angry, Calm

Tweet classification containing all moods

This section contains the figures presenting the predicted moods for the two datasets of COVID19-related tweets from both India as well all over the World. For plotting these figures, all of the 25 possible moods have been considered (Figs. 12, 13, 14, 15, 16, 17, 18). These results have been included in appendix for the sake of completeness of the experimental results presented.

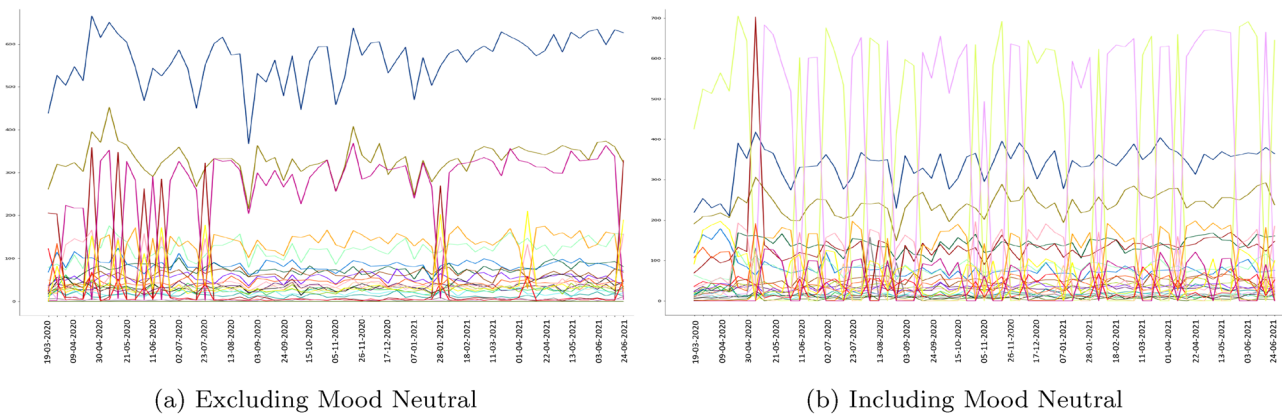**Fig. 12** Legends for Figs. 13 , 14, 15, 16, 17 and 18

(a) Excluding Mood Neutral

(b) Including Mood Neutral

Fig. 13 Tweets from Over the world classified using Random Forest. Data points are 7 days apart on *x*-axis

(a) Excluding Mood Neutral

(b) Including Mood Neutral

Fig. 14 Tweets from India classified using Random Forest

(a) Excluding Mood Neutral

(b) Including Mood Neutral
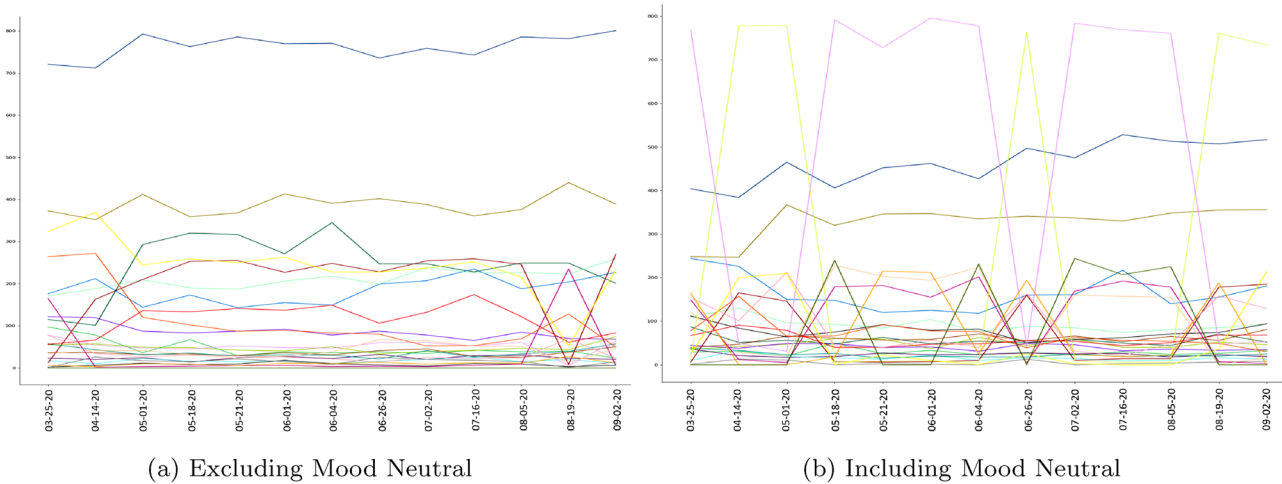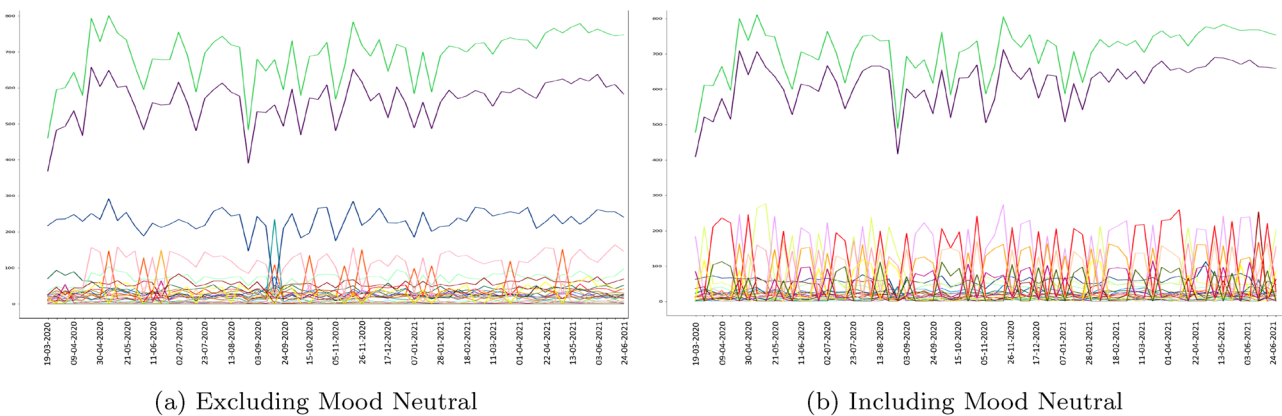
Fig. 15 Tweets from Over the world classified using Decision Tree. Data points are 7 days apart on *x*-axis

(a) Excluding Mood Neutral

(b) Including Mood Neutral

Fig. 16 Tweets from India classified using Decision Tree
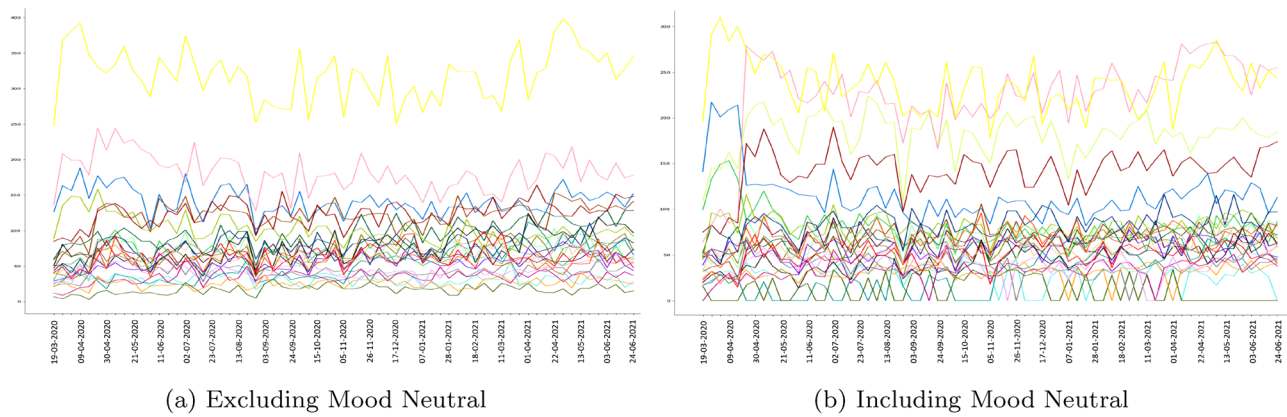

(a) Excluding Mood Neutral

(b) Including Mood Neutral
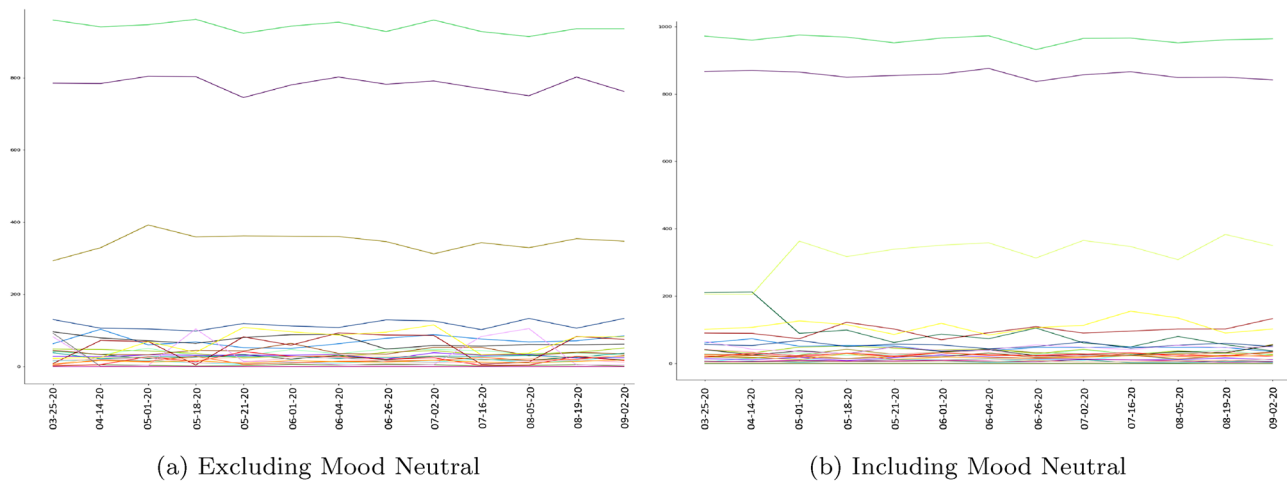
Fig. 17 Tweets from Over the world classified using Complement Naive Bayes. Data points are 7 days apart on *x*-axis
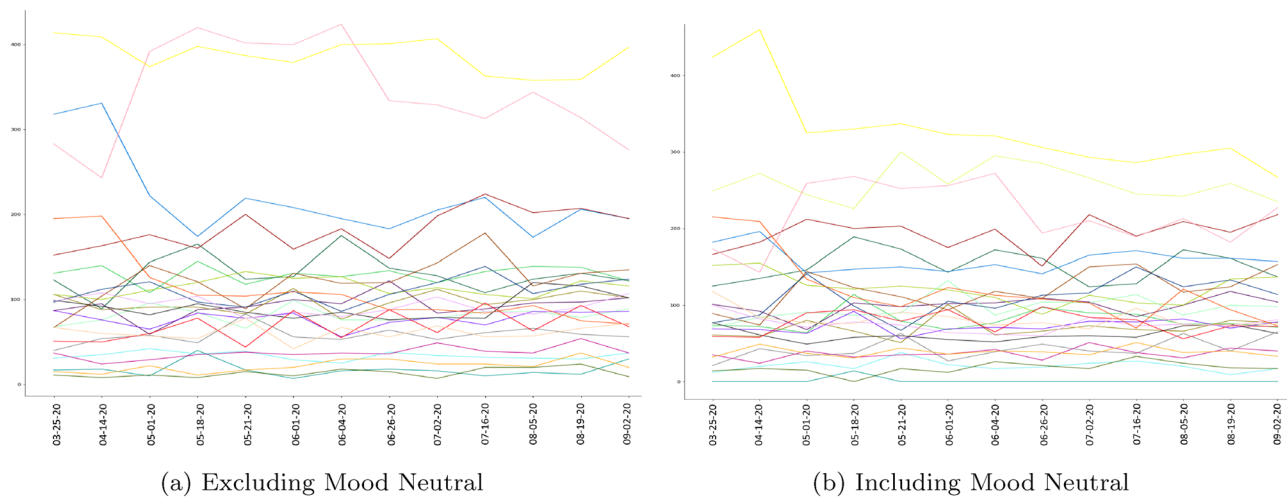

(a) Excluding Mood Neutral

(b) Including Mood Neutral

Fig. 18 Tweets from India classified using Complement Naive Bayes

App screenshots

The following pictures are screenshots from the CitizenSense application depicting some of the real posts and moods posted by users (Figs. 19, 20).
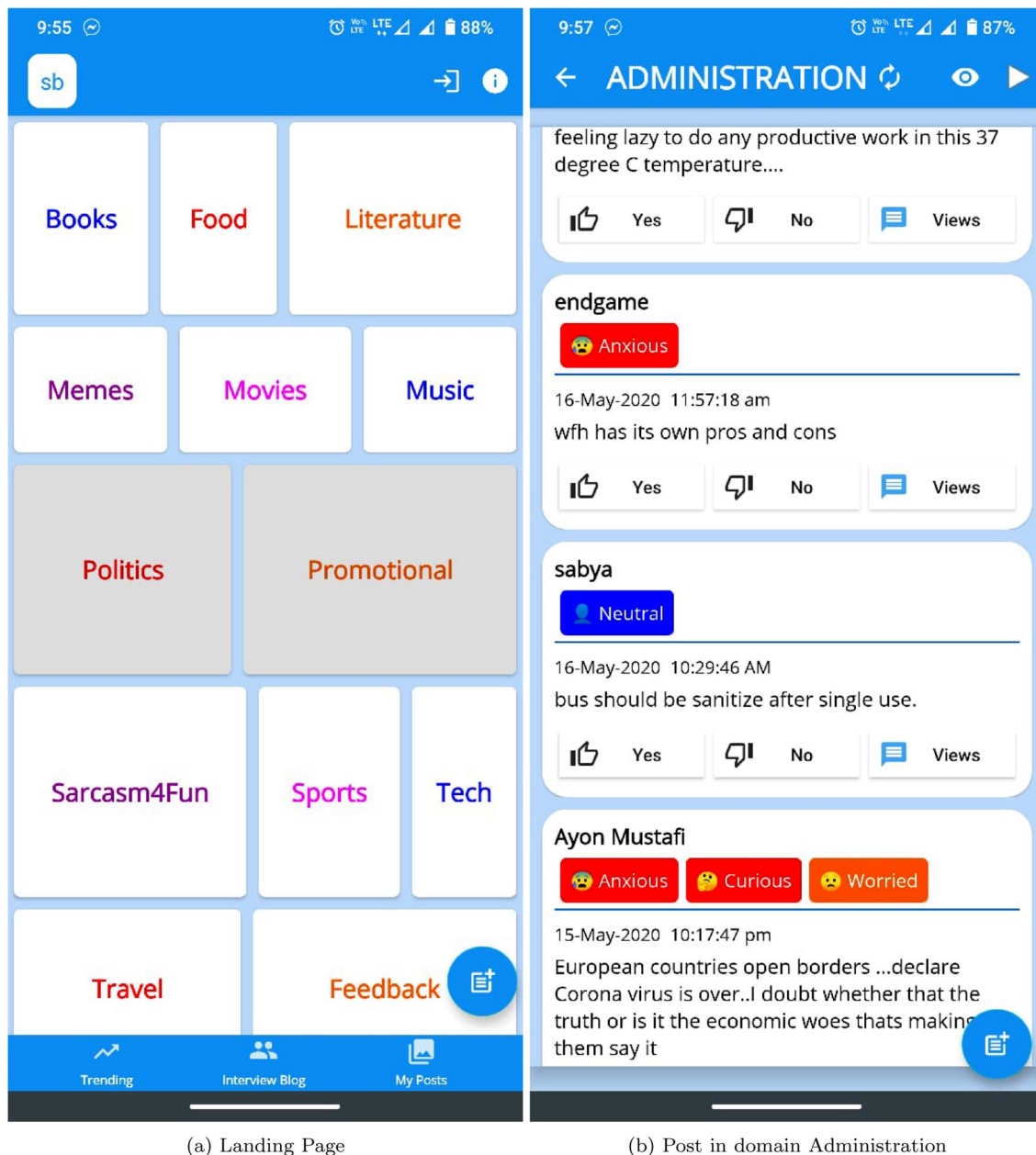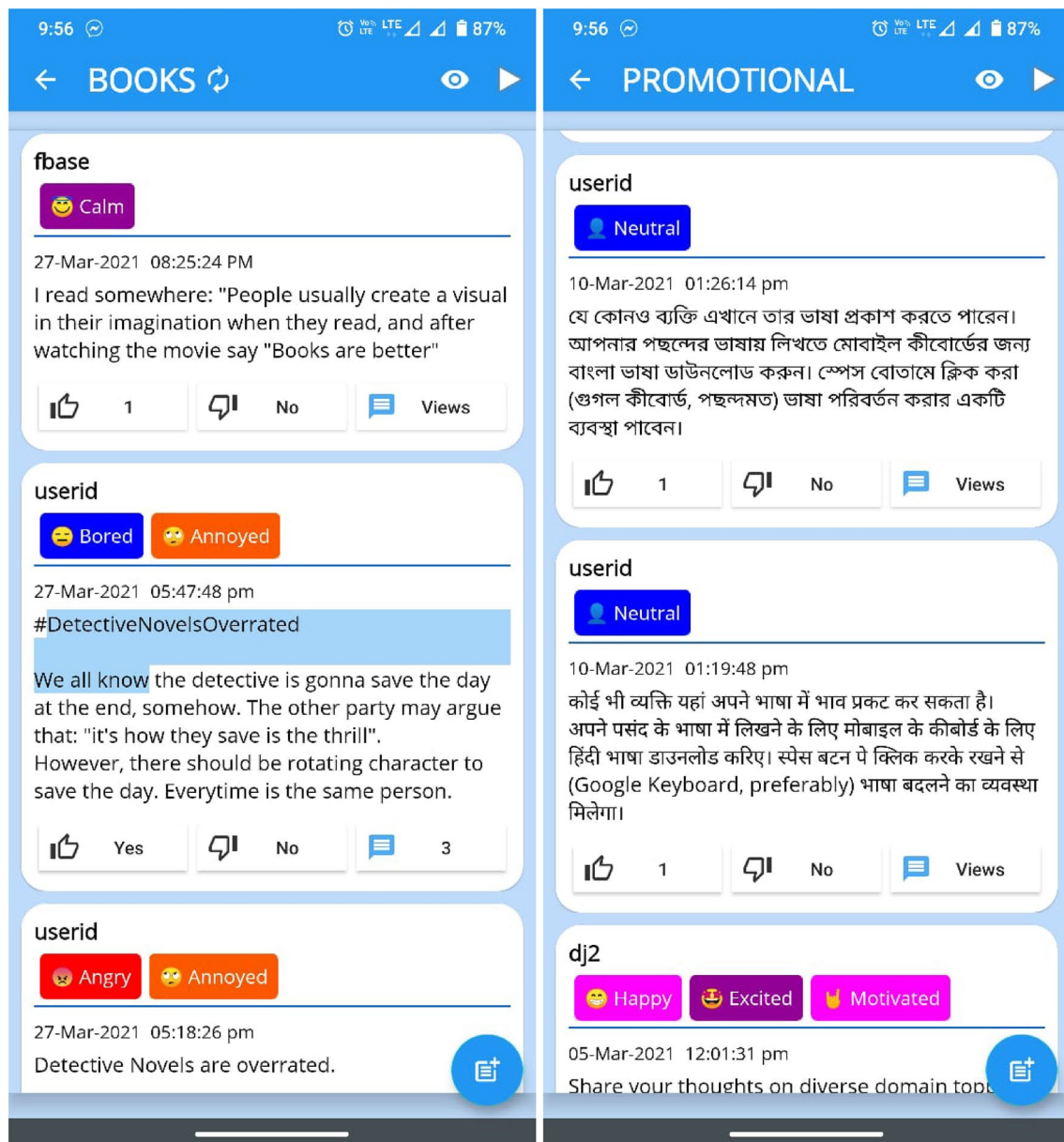

(a) Landing Page   (b) Post in domain Administration

Fig. 19 **a** Landing Page **b** Post in domain administration

(a) Post in domain Books

(b) Post in domain Promotional

Fig. 20 **a** Post in domain books **b** Post in domain promotional

Stratified sampling results

The following are the results for accuracy, recall, precision and F1 score for each of the classifiers for each of the moods where the training data is prepared using stratified sampling (Fig. 21, Table 12 and 13).
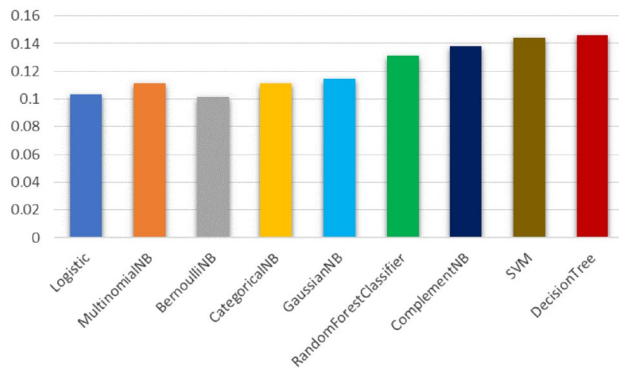

Fig. 21 Cumulative accuracy of classifiers

Table 12 Validation—precision (*p*), recall (*r*) and F1 score for moods using logistic regression, multinomial Naive Bayes, Bernoulli Naive Bayes, and Complement Naive Bayes

	Logistic			MultinomialNB			BernoulliNB			ComplementNB		
	r	p	F1	r	p	F1	r	p	F1	r	p	F1
Angry	0.53	0.50	0.52	0.64	0.37	0.47	0.42	0.70	0.53	0.50	0.38	0.43
Excited	0.47	0.28	0.35	0.40	0.47	0.43	0.10	0.56	0.17	0.38	0.37	0.38
Calm	0.32	0.38	0.35	0.47	0.35	0.40	0.88	0.15	0.26	0.38	0.31	0.34
Curious	0.59	0.23	0.33	0.33	0.78	0.46	0.14	1.00	0.24	0.49	0.42	0.45
Relaxed	0.46	0.25	0.32	0.37	0.40	0.38	0.09	0.80	0.16	0.40	0.40	0.40
Worried	0.39	0.45	0.42	0.47	0.44	0.45	0.63	0.16	0.25	0.52	0.45	0.48
Annoyed	0.50	0.36	0.42	0.43	0.43	0.43	0.11	1.00	0.19	0.56	0.48	0.52
Scared	0.73	0.50	0.59	0.76	0.39	0.51	0.63	0.41	0.49	0.80	0.34	0.48
Anxious	0.60	0.66	0.63	0.58	0.44	0.50	0.46	0.30	0.36	0.50	0.42	0.45
Motivated	0.28	0.26	0.27	0.50	0.22	0.31	0.07	1.00	0.14	0.44	0.36	0.40
Neutral	0.23	0.33	0.27	0.33	0.44	0.38	0.14	0.50	0.21	0.31	0.25	0.28
Sad	0.49	0.30	0.37	0.83	0.14	0.24	0.41	0.74	0.53	0.53	0.37	0.44
Greedy	0.86	0.37	0.51	0.88	0.15	0.26	0.86	0.44	0.58	0.87	0.25	0.38
Stressed	0.37	0.42	0.40	0.47	0.43	0.45	0.18	0.88	0.30	0.55	0.47	0.51
Bored	0.44	0.27	0.33	0.37	0.63	0.47	0.72	0.35	0.47	0.55	0.33	0.41
Happy	0.41	0.49	0.45	0.22	0.43	0.29	0.10	0.50	0.16	0.27	0.37	0.31
Confused	0.82	0.19	0.31	0.49	0.49	0.49	0.02	1.00	0.04	0.49	0.46	0.47
Exhausted	0.58	0.44	0.50	0.55	0.47	0.51	0.94	0.14	0.24	0.44	0.39	0.41
Hopeless	0.61	0.27	0.37	0.67	0.37	0.47	0.21	0.89	0.33	0.69	0.35	0.46
Sorry	0.93	0.81	0.87	1.00	0.38	0.55	0.86	0.75	0.80	0.93	0.41	0.57
Mad	0.58	0.38	0.45	0.60	0.35	0.44	0.11	1.00	0.20	0.77	0.35	0.48
Nervous	0.77	0.49	0.60	0.82	0.32	0.46	0.52	0.46	0.49	0.82	0.33	0.47
Grateful	0.55	0.44	0.49	0.41	0.81	0.54	0.70	0.31	0.43	0.59	0.35	0.44
Proud	0.76	0.37	0.50	0.84	0.33	0.47	0.37	1.00	0.54	0.88	0.27	0.42
Lonely	0.82	0.24	0.37	0.92	0.16	0.27	0.27	1.00	0.43	1.00	0.18	0.31

Table 13 Prediction—precision (*p*), recall (*r*) and F1 score for moods using Categorical Naive Bayes, Gaussian Naive Bayes, Random Forest, support vector machine (SVM), and Decision Tree

	CategoricalNB			GaussianNB			RandomForest-Classifier			SVM			DecisionTree		
	r	*p*	F1	*r*	p	F1	*r*	*p*	F1	*r*	*p*	F1	*r*	*p*	F1
Angry	0.06	1.00	0.11	0.88	0.08	0.14	0.63	0.35	0.45	0.56	0.80	0.66	1.00	0.08	0.14
Excited	0.98	0.11	0.20	0.26	0.43	0.32	0.58	0.37	0.46	0.60	0.29	0.39	0.31	0.61	0.41
Calm	0.03	1.00	0.06	0.32	0.49	0.38	0.42	0.38	0.39	0.44	0.37	0.40	0.16	0.45	0.23
Curious	0.00	0.00	0.00	0.44	0.42	0.43	0.63	0.35	0.45	0.63	0.27	0.38	0.49	0.56	0.52
Relaxed	0.00	0.00	0.00	0.36	0.34	0.35	0.49	0.27	0.35	0.64	0.37	0.47	0.27	0.46	0.34
Worried	0.02	0.50	0.03	0.22	0.48	0.31	0.40	0.43	0.41	0.47	0.33	0.38	0.38	0.75	0.51
Annoyed	0.00	0.00	0.00	0.69	0.10	0.18	0.48	0.37	0.42	0.50	0.43	0.46	0.82	0.14	0.24
Scared	0.04	1.00	0.07	0.58	0.63	0.60	0.56	0.52	0.54	0.63	0.86	0.73	0.54	0.82	0.65
Anxious	0.02	1.00	0.05	0.46	0.40	0.43	0.55	0.47	0.51	0.45	0.70	0.55	0.61	0.43	0.51
Motivated	0.00	0.00	0.00	0.27	0.50	0.35	0.61	0.17	0.26	0.52	0.21	0.30	0.35	0.35	0.35
Neutral	0.00	0.00	0.00	0.05	0.17	0.07	0.28	0.25	0.26	0.37	0.21	0.27	0.14	0.38	0.20
Sad	0.00	0.00	0.00	0.43	0.50	0.46	0.64	0.51	0.57	0.63	0.49	0.55	0.66	0.61	0.63
Greedy	0.00	0.00	0.00	0.82	0.50	0.62	0.83	0.48	0.60	0.77	0.59	0.67	0.65	0.63	0.64
Stressed	0.00	0.00	0.00	0.51	0.45	0.48	0.43	0.38	0.40	0.59	0.52	0.55	0.35	0.54	0.43
Bored	0.00	0.00	0.00	0.46	0.35	0.40	0.50	0.38	0.43	0.51	0.47	0.49	0.45	0.68	0.54
Happy	0.00	0.00	0.00	0.24	0.65	0.35	0.33	0.35	0.34	0.69	0.29	0.41	0.16	0.50	0.25
Confused	0.02	1.00	0.05	0.29	0.56	0.38	0.46	0.47	0.47	0.58	0.38	0.46	0.50	0.79	0.61
Exhausted	0.98	0.10	0.18	0.31	0.44	0.36	0.47	0.35	0.40	0.47	0.40	0.43	0.29	0.55	0.38
Hopeless	0.00	0.00	0.00	0.39	0.57	0.46	0.54	0.54	0.54	0.57	0.40	0.47	0.54	0.74	0.63
Sorry	0.07	1.00	0.13	1.00	0.68	0.81	1.00	0.57	0.72	0.86	1.00	0.92	0.93	0.88	0.90
Mad	0.00	0.00	0.00	0.64	0.53	0.58	0.77	0.47	0.59	0.68	0.79	0.73	0.68	0.65	0.67
Nervous	0.14	1.00	0.25	0.71	0.59	0.64	0.78	0.46	0.58	0.87	0.87	0.87	0.74	0.65	0.69
Grateful	0.03	1.00	0.06	0.50	0.59	0.54	0.54	0.44	0.48	0.61	0.57	0.59	0.50	0.73	0.59
Proud	0.95	0.04	0.07	0.84	0.84	0.84	0.83	0.56	0.67	0.73	1.00	0.85	0.85	0.81	0.83
Lonely	0.10	0.50	0.17	0.92	0.44	0.59	1.00	0.34	0.51	0.90	0.64	0.75	0.80	0.53	0.64

Code Availability The code used for this work can be made available on request.

Data Availability Most of the data used as a part of this work is open source, and the sources have been cited. The novel data for application CitizenSense can be made available on request.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

References

Agrawal A, An A (2012) Unsupervised emotion detection from text using semantic and syntactic relations. In: Proceedings of the The 2012 IEEE/WIC/ACM international joint conferences on web intelligence and intelligent agent technology - Volume 01, WI-IAT '12, pp 346–353, USA. IEEE Computer Society

Aldunate N, Villena-González M, Rojas-Thomas F, López V, Bosman CA (2018) Mood detection in ambiguous messages: the interaction between text and emoticons. Front Psychol 9:423

An Y, Sun S, Wang S (2017) Naive Bayes classifiers for music emotion classification based on lyrics. In: 2017 IEEE/ACIS 16th international conference on computer and information science (ICIS), pp 635–638

Anagnostopoulos C-N, Iliou T, Giannoukos I (2012) Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011. Artif Intell Rev 43:02

Bandhakavi A, Wiratunga N, Padmanabhan D, Massie S (2017) Lexicon based feature extraction for emotion text classification. Pattern Recognit Lett, 93:133 – 142. Pattern recognition techniques in data mining

Bhatt A, Patel A, Chheda H, Gawande K (2015) Amazon review classification and sentiment analysis. Int J Comput Sci Inf Technol 6(6):5107–5110

Binali H, Wu C, Potdar V (2010) Computational approaches for emotion detection in text. In: 4th IEEE international conference on digital ecosystems and technologies, pp 172–177

Bodendorf F, Kaiser C (2009) Detecting opinion leaders and trends in online social networks. In: Proceedings of the 2nd ACM workshop on Social web search and mining, pp 65–68

Dini L, Bittar A (2016) Emotion analysis on twitter: the hidden challenge. In: Proceed Tenth Int Conf Lang Resour Eval (LREC'16), pp 3953–3958

Hasan M, Rundensteiner E, Agu E (2018) Automatic emotion detection in text streams by analyzing twitter data. Int J Data Sci Anal 7(1):35–51

Hasan M, Rundensteiner E, Agu E (2019) Automatic emotion detection in text streams by analyzing twitter data. Int J Data Sci Anal 7(1):35–51

https://play.google.com/store/apps/details?id=com.citizensense.android.campuscare

Hussein A, Al Kafri M, Abonamah AA, Tariq MU (2020) Mood detection based on Arabic text documents using machine learning methods. Intl J 9(4):4224–4336

Kang K, Yoon C, Kim EY (2016) Identifying depressive users in twitter using multimodal analysis. In: 2016 international conference on big data and smart computing (BigComp), pp 231–238. IEEE

Kharde V, Sonawane P, et al (2016) Sentiment analysis of twitter data: a survey of techniques. arXiv preprint arXiv:1601.06971

Kulkarni SS, Reddy NP, Hariharan SI (2009) Facial expression (mood) recognition from facial images using committee neural networks. Biomed Eng 8(1):1–12

Lamsal R (2020) Coronavirus (covid-19) tweets dataset

Lamsal R (2020) Tweets originating from india during covid-19 lockdowns

Li X, Pang J, Mo B, Rao Y (2016) Hybrid neural networks for social emotion detection over short text. In: 2016 international joint conference on neural networks (IJCNN), pp 537–544

Li X, Pang J, Mo B, Rao Y (2016) Hybrid neural networks for social emotion detection over short text. In: 2016 International joint conference on neural networks (IJCNN), pp 537–544. IEEE

Mohammad S (2018) Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In: Proceedings of the 56th annual meeting of the association for computational linguistics (Vol. 1: Long Papers), pp 174–184

Mohammad SM, Bravo-Marquez F (2017) Emotion intensities in tweets. arXiv preprint arXiv:1708.03696

Ramaswamy S, Mathews R, Rao K, Beaufays F (2019) Federated learning for emoji prediction in a mobile keyboard. arXiv preprint arXiv:1906.04329

Sailunaz K, Dhaliwal M, Rokne J, Alhajj R (2018) Emotion detection from text and speech: a survey. Soc Netw Anal Min 8(1):28

Scott S, Matwin S (1999) Feature engineering for text classification. In: ICML, vol 99, pp 379–388. Citeseer

Sen A, Sinha M, Mannarswamy S, Roy S (2017) Multi-task representation learning for enhanced emotion categorization in short text. In: Pacific-Asia conference on knowledge discovery and data mining, pp 324–336. Springer

Summa A, Resch B, Strube M (2016) Microblog emotion classification by computing similarity in text, time, and space. In: Proceedings of the workshop on computational modeling of people's opinions, personality, and emotions in social media (PEOPLES), pp 153–162

Urmi NP, Ahmed NU, Sifat MHR, Islam S, Mohammad Mahdee JAS (2020) Banglamusicmood: A music mood classifier from bangla music lyrics. In: International conference on mobile computing and sustainable informatics, pp 673–681. Springer

Xu H, Quang Phan T , Tan B (2013) How does online social network change my mood? An empirical study of depression contagion on social network sites using text-mining

Yong Lu, Luo Xin, Polgar Michael, Cao Yuanyuan (2010) Social network analysis of a criminal hacker community. J Comput Inf Syst 51(2):31–41