An Exhaustive Sentiment and Emotion Analysis of COVID-19 Tweets Using Machine Learning, Ensemble Learning and Deep Learning Techniques

Jasleen Kaur, Smit Patel, Meet Vasani, and Jatinderkumar R. Saini

Abstract COVID-19 has been generating new variations one after the other and there is no end to it. Even though vaccines are out, the cases are skyrocketing after each day while the number of deaths has increased simultaneously. In these crucial times, it is necessary to build a system which can aid in making the situation controlled by taking the necessary actions. There are number of ways available to deal with this situation and it is very much essential to highlight those different steps which can help not only in the advancement of technology but also will replenish the goal of thinking different when any pandemic strikes again, if at all, in the future. The main purpose to carry out this research is to exhaustively understand the 3 sentiments (positive, negative and neutral) as well as 11 emotions (Optimistic, Thankful, Empathetic, Pessimistic, Anxious, Sad, Annoyed, Denial, Surprise, Official report, Joking) of public towards COVID-19 pandemic. 5000 COVID-19 related tweets were collected from Twitter and different perspectives such as government policies, safety measures, COVID-19 symptoms and precautionary measures were considered for sentiment analysis as well as emotion detection task which was performed using 12 different models. These models were categorized as baseline models, ensemble learning models and deep learning models. Results revealed that ensemble learning models outperformed baseline and deep learning models for sentiment analysis task. Highest accuracy 60.1% was reported by Gradient boosting algorithm. For emotion analysis task, baseline category performed better as compared to ensemble and deep learning models. Finally, Multinomial Naïve Bayes was reported as the winning algorithm.

Keywords COVID-19 · Deep learning · Emotion detection · Ensemble learning · Machine learning · Sentiment analysis

J. Kaur · S. Patel · M. Vasani
School of Engineering, PP Savani University, Surat, India

J. R. Saini (✉)
Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Pune, India
e-mail: saini_expert@yahoo.com

V. Goar et al. (eds.), *Advances in Information Communication Technology and Computing*, Lecture Notes in Networks and Systems 628,
https://doi.org/10.1007/978-981-19-9888-1_36

1 Introduction

The coronavirus pandemic gave horrific scenes to the crowd around world in the years 2020 and 2021. As there were no vaccines or medicines, which can cure the symptoms of the virus, the different Governments took various measures which were not formal in their own ways but kept its place [1, 2]. Various precautions included quarantine, lockdown, self-isolation, safe distance, and many more which by some extent reduced the havoc which shook the whole world.

This was one part of the whole situation going around in the world, well the other half was occupied by activists, political leaders, the people itself, media and many more. But the major part was occupied by social media. Nowadays, Social Media has become a key to express positivity around, but it also leads to negative impact when the information is quite accurate or satisfy the human mind and so does it happen in the COVID -19 pandemic, where number of news and rumors were communicated across the Social Media platform, making people excited and happy when they see a tweet or a post which relieves them from the pain of virus itself and also at the same time making people sad and more panicked when they see something which is unimaginable. So, there was a need where posts, tweets etc. get a proper segregation which can define whether that particular post will make a healthy impact on the readers' mind or will give a shocking impact.

The main purpose of this research work is to perform emotion detection and sentiment analysis of tweets over COVID-19. To carry out this work, various machine learning, ensemble and deep learning methods are used for extracting emotion based sentiments from tweets. The paper is organized as follows. Section 2 presents related work carried out in proposed direction. Complete methodology followed for implementation of this work is presented in Sect. 3. Detailed results and analysis are presented in Sect. 4 followed by conclusion in Sect. 5.

2 Previous Work

This section presents previous reported work carried out in this direction. Table 1 presents approach and feature based analysis of various works carried related to sentiment analysis of COVID-19.

From Table 1, it can be observed that number of different techniques (including supervised and unsupervised) were experimented to identify sentiments related to COVID-19 from different social media platforms. Figure 1 depicts the pictorial distribution for different approaches used. From machine learning (ML) area, prominent algorithm usage is for Support Vector Machine (SVM), Naïve Bayes (NB) and Logistic regression (LR) [3–5]. Deep Learning (DL) techniques such as Long short term memory (LSTM), Bi-directional Encoder Representations from Transformers (BERT) and Bi-directional Long Short Term Memory (Bi-LSTM) were used by many researchers [6–9]. Other techniques used by different researchers include

Table 1 Approach and feature based analysis for COVID-19 sentiment analysis

Year	Approach used	Feature used	Performance/outcome	Ref.
2021	MLP	Bag of word	83%	[6]
2021	GSOM	Word2vec	Negative emotions	[10]
2021	LDA	–		[11]
2021	LSTM, Bi-LSTM	Bag of word	90.59%, 90.83%	[6]
2021	Bernoulli NB, SVM, LR	VADER based sentiment score	85.95%	[3]
2021	LSTM, Bi-LSTM, BERT	Bag of word	49%, 58%	[7]
2021	LDA	Bag of word	73% (RNN)	[12]
2021	Lexicon based Kruskall-Wallis test	Bag of word	p-value < 0.0001	[13]
2021	Lexicon based	Bag of word	Positive sentiment	[14]
2021	Bi-LSTM	GloVe	89.51%	[15]
2021	LDA	Bag of Word	Vaccine hesitancy among users	[16]
2021	BERT	–	87%	[8]
2021	Clustered based	–	98%	[17]
2021	NB	Bag of word	81.77%	[4]
2020	LDA	Bag of word	Positive emotions	[18]
2020	LR	–	78.5%	[5]

unsupervised learning (US) techniques, Latent Dirichlet Allocation (LDA), Multi-layer Perceptron (MLP), Growing Self-organizing Map (GSOM) and lexicon-based (LB) Natural Language Processing (NLP). These techniques were implemented for extraction of emotions and sentiments from COVID-19 text on different social media platforms. Figure 2 presents the average based performance analysis of existing COVID-19 sentiment and emotion detection analysis works.

Dataset based analysis was provided in Table 2. Sentiment and Emotion detection analysis for COVID-19 related text was carried in different perspective such as false

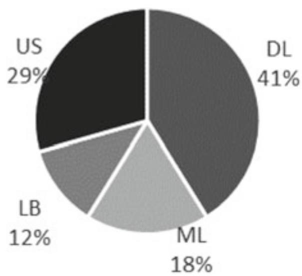


Fig. 1 Different approaches used for COVID-19 sentiment analysis

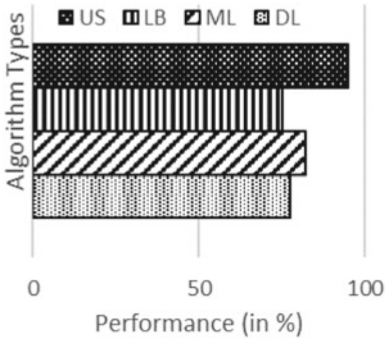


Fig. 2 Performance analysis of existing COVID-19 sentiment analysis works

news during COVID-19, COVID-19 related awareness, COVID-19 vaccination opinions, COVID-19 and political perspective, public response to COVID-19 and many more. Time span considered for COVID-19 analysis is March 2020 to May 2021. Figure 3 depicts that much of sentiment analysis work was carried during first wave of COVID-19. Language of majority of tweets is English. From reviewed literature, it can be concluded that for sentiment analysis task, main class labels used are positive, negative and neutral whereas for emotion detection task, main class labels are fear, sadness, anger, disgust and optimistic.

3 Methodology

Architecture of proposed methodology is depicted in Fig. 4. Proposed system consists of two main phases: phase 1 and phase 2. The detail description of phase 1 and phase 2 is presented in Fig. 4.

3.1 Phase 1

It consists of the following sub phases.

3.1.1 Data Collection and Understanding the Dataset

For this research work, we have utilized admission dataset from Kaggle [5]. Dataset comprises of 5000 tweets which were further divided into categories and sub categories. For sentiment analysis of COVID-19 related tweets, tweets were bifurcated into positive, negative and neutral tweets. For emotion detection related to COVID-19 pandemic, 11 different labels were selected. This sub categorization includes: Optimistic (0), Thankful (1), Empathetic (2), Pessimistic (3), Anxious (4), Sad (5),

Table 2 COVID-19 dataset analysis for sentiment analysis task

Ref.	Description	Source	Volume	Duration	Class count	Language	Location
[12]	Vaccination	Twitter	125,906	December 20, 2020 to July 21, 2021	3	English	Across world
[13]	Leader's response on COVID-19	Twitter	15,848	January 1, 2020 to December 21, 2020	2	English	United Kingdom, Canada, New Zealand, America
[14]	Sentiment analysis	Twitter	150,000	March 2020 to September 2020	5	English	India
[15]	Public sentiments	Twitter	4511	February 1, 2020 to August 31, 2020	3	English	Singapore
[16]	Sentiments about vaccination	Twitter	701 891	December 1 2020 to March 31, 2021	2	English	Across the World
[10]	Sentiments about vaccination	Twitter	4 million	January 2020 to January 2021	3	English	Across the world
[8]	Emotion detection about COVID-19	Twitter	5,60,14,158	March 5, 2020 to December 31, 2020	10	English	USA
[17]	Public Opinion about COVID-19 vaccination	Reddit	18,000	December 1, 2020 to May 15, 2021	2	English	Across the world
[3]	Emotion Detection	Twitter	73,000	January 2020 to September 2020	8	English	Australia

(continued)

Table 2 (continued)

Ref.	Description	Source	Volume	Duration	Class count	Language	Location
[9]	Emotion detection about COVID-19 vaccination	Twitter	928,402	–	4	English Turkish	USA, UK, Canada, Turkey, France, Germany, Spain and Italy
[18]	Sentiment analysis	Twitter	0.3 million	Before March 20, 2020	3	English	Across the world
[6]	Sentiment analysis	Twitter	70,000	–	3	English	Nepal and Italy
[19]	Public response to the COVID-19 pandemic	Twitter	4 million	March 7, 2020 to April 21, 2020	5	English	Across the world
[7]	Sentiments Relating to COVID-19 Vaccination	Twitter	–	January 2020 to October 2020	8	English	Australia
[11]	Sentiment related to COVID-19 vaccination	Twitter	993	March 1, 2021 to March 31, 2021	3	English	Philippines
[4]	Sentiment analysis	Twitter	50,000	January 2020 to July 2020	2	English	Across the world

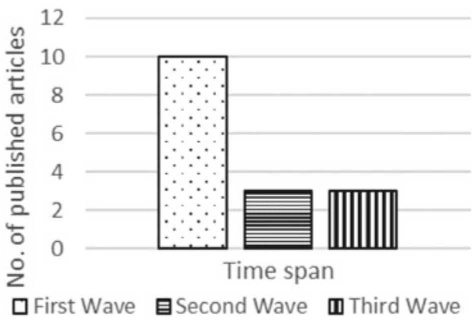


Fig. 3 COVID-19 sentiment analysis time span

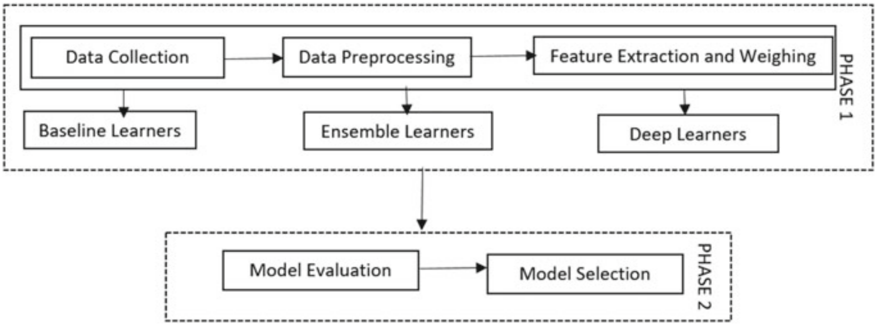


Fig. 4 Architecture of COVID-19 sentiment and emotion detection system

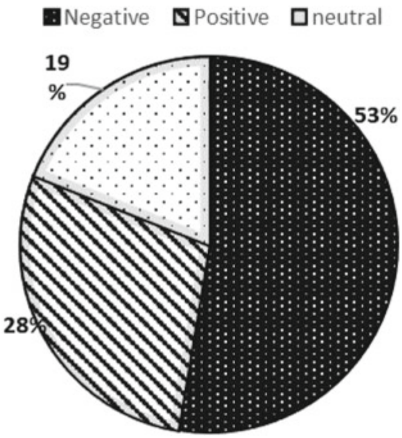
Annoyed (6), Denial (7), Surprise (8), Official report (9), Joking (10). Statistical analysis of the dataset is provided in Table 3.

Figures 5 and 6 show the distribution of dataset (in categories and sub categories). Basic experimental analysis was performed to understand human emotions in 3 polarities, i.e., positive, negative, and neutral; our findings showed that 28% of people were positive, 52% were negative, and 19% were neutral, in response to COVID-19 worldwide. Emotion based classes distribution was presented in Fig. 6. Out of 10 emotion labels, prominent distribution of tweets was present in optimistic (23%), annoyed (17%), sad (13%), anxious (11%).

Table 3 Dataset description

Dataset	Total tweets	Total words	Average tweet length
COVID-19	5000	51,388	10.27

Fig. 5 Sentiment class distribution in dataset



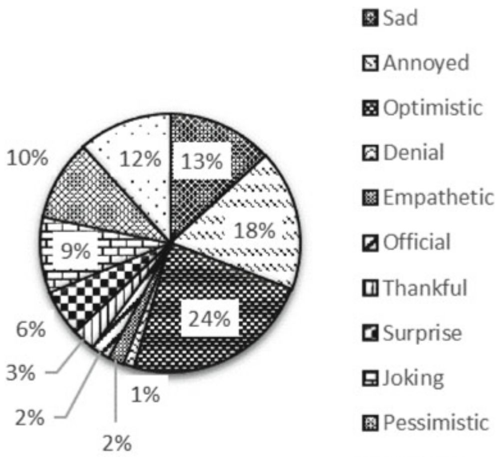


Fig. 6 Emotion class distribution in dataset

(a) Negative (b) Positive (c) Neutral

Fig. 7 Word Cloud for 3 sentiment analysis class labels

For better understanding of dataset, word analysis was carried out. Top words in each category of sentiment (positive, negative, neutral) were presented in Fig. 7. Table 4 presents the top-5 words present in each.

3.1.2 Data Pre-processing

All tweets were passed through various pre-processing phases:

1. Removing Numbers, Special Characters and Punctuations

Punctuation marks, numbers, and special characters are not helpful in analyzing emotions. It is best to remove them from the text. Here we will replace everything except letters with spaces.

Table 4 Top 5 words in each emotion class label

Sub category	Top 5 words
Optimistic	Coronavirus, Covid, great, joke, doctors
Annoyed	Corona, pencils, coronavirus, Covid, China
Pessimistic	Corona, virus, world, American, temperatures
Surprise	Coronavirus, death, cases, new, India
Sad	Coronavirus, life, corona, bat, government
Joking	Corona, coronavirus, April, virus, neighbours
Anxious	Corona, enough, medication, critical, effects
Thankful	Health, therapy, perfect, greatest, books
Empathetic	Please, corona, coronavirus, joke, end
Denial	Diagnostics, China, deaths, long, bioweapon
Official	God, internet, virus, natural, elaborate

2. Stopwords Removal

In NLP work stopwords (very common words e.g., that, are, have) do not make sense in reading because they are not connected with emotions. Removing them therefore saves integration and increases the accuracy of the model.

3. Stemming using Porter Stemmer

Stemming is used to remove the suffixes such as ('-ing', '-ly', '-es', '-s', etc.) to get a root word of some particular word specified. We implemented Porter Stemmer in our work. We have used five step process, all with its own rules. Porter Stemmer is renowned because of its easy-to-use behavior, speed and efficiency. The outcome will get us a word in its root form.

4. Label Encoding of target variables

This is an encoding which converts the categorical values in integer values in between the range of 0 and the number of classes minus 1. If suppose, we have 5 distinct categorical classes, then the conversion would be (0, 1, 2, 3, 4).

3.1.3 Feature Extraction and Feature Weighing

After pre-processing of data, 'Bag of Word' model is used for feature extraction and vector space representation was created for entire data. Term frequency (TF) and Term-frequency inverse document frequency (TF-IDF) is used for feature weighing.

3.1.4 Model Training

In total, 12 models were trained and tested on this dataset. Based on their type, these models were divided into three categories: Baseline Learners (BL), Ensemble

Learners (EL) and Deep Learners (DL). Baseline learners consists of Logistic Regression (LR), K- Nearest Neighbour (KNN), Support Vector Machine (SVM), Multinomial Naïve Bayes (MNB) and Decision Tree (DT). LR, NB (& it's variation), SVM are statistical models in nature. LR is a way of modeling probability of a discrete outcome given an input variable [3, 5]. NB is based on conditional probability and Bayes theorem [3, 4]. SVM perform classification by finding a hyperplane that distinctly classifies the data points [3, 4]. Ensemble Learners consists of Random Forest (RF), XG Boost (XGB), Bagging (BG) and Gradient Boosting (GB) [20]. RF operated by constructing multitude of decision tree. XGB uses gradient boosting technique to generate boosted tree with enhanced performance. BG aggregates the performance of several weak models. GB tries to minimize the loss function by adding weak learners using gradient descent. Deep learners consist of Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) [6, 7. ANN is nonlinear statistical model which exhibits the complex relationship between input and output. CNN is a class of deep neural network which consists of an input layer, an output layer and numerous hidden layers. LSTM is one type of recurrent neural network that records different cell state to perform the classification.

3.1.5 Phase 2

Performance evaluation is carried out using Accuracy, Precision, Recall and F1-measure [21].

4 Results and Analysis

This section provides the result and analysis of the application of 12 algorithms on two feature weighing criteria (TF, and TF-IDF) and on sentiment analysis as well as emotion detection tasks.

4.1 Result and Analysis on Sentiment Analysis

4.1.1 Sentiment Analysis Results Using TF

From Table 5, it can be observed that, with accuracy of 59.1%, precision of 64.9%, recall of 46.3% and F1-score of 47.3%, SVM performed better as compared to other baseline algorithms followed by Multinomial Naïve Bayes. Among ensemble learning methods, gradient boosting turns out to be the best with accuracy, precision, recall and F1-score of 60.1%, 63.1%, 48.2% and 49.5%, respectively. It can be observed that with highest accuracy, precision, recall and F1-score (34.9%, 54.0%,

Table 5 Results of algorithms using term frequency as feature weighing

Type	Model	Accuracy	Precision	Recall	F1-score
Baseline learners (BL)	LR	59.3	57.0	51.4	52.9
	KNN	52.8	49.7	39.3	36.7
	SVM	59.1	64.9	46.3	47.3
	MNB	59.0	56.3	50.6	52.0
	DT	52.3	46.9	46.0	46.3
	Average	56.5	55.0	46.7	47.0
Ensemble learners (EL)	RF	57.5	56.0	46.5	47.5
	XGB	59.3	58.3	47.8	48.9
	BG	58.4	56.6	48.2	49.4
	GB	60.1	63.1	48.2	49.5
	Average	58.8	58.5	47.7	48.8
Deep learners (DL)	ANN	20.9	49.4	98.0	64.6
	CNN	34.9	54.0	81.5	64.3
	LSTM	19.8	47.9	99.6	64.8
	Average	25.2	50.4	93.0	64.6

81.5% and 64.3%, respectively), CNN turns out to be the best among deep learning methods.

4.1.2 Sentiment Analysis Results Using TF-IDF

From the Table 6, it can be observed that, with an accuracy of 59.3%, precision of 66.5%, recall of 46.2% and F1-score of 47.1%, SVM performed better compared to other baseline learners followed by Logistic Regression. From ensemble learning category, gradient boosting has become the best among the ensemble learners. Accuracy, precision, recall and F1-score in gradient boosting were reported to be 58.8%, 58.9%, 47.1% and 48.2%, respectively. CNN turns out to be the best among the deep learners.

Figure 8 indicates that ensemble learners performed better as compared to other ones. Performance of TF and TF-IDF is approximately equal for sentiment analysis task. **From review of existing state-of-art research carried out in this direction (as represented in Table 1), ensemble learning techniques have never been applied for sentiment as well as emotion detection work**. Deep learners were not suitable for sentiment analysis task. Analysis based on other performance metrics (Precision, Recall, F1-Score) are presented in Fig. 9.

Table 6 Results of algorithms using TF-IDF as feature weighing

Type	Model	Accuracy	Precision	Recall	F1-score
Baseline learners	LR	59.2	63.8	46.8	47.9
	KNN	58.0	59.2	46.2	47.2
	SVM	59.3	66.5	46.2	47.1
	MNB	57.5	74.9	41.1	38.2
	DT	50.8	45.2	44.3	44.5
	Average	57.0	61.9	44.9	45.0
Ensemble learners	RF	57.6	55.5	44.8	44.9
	XGB	57.3	62.4	43.1	42.6
	BG	56.9	55.4	46.8	48.0
	GB	58.8	58.9	47.1	48.2
	Average	57.7	58.1	45.5	45.9
Deep learners	ANN	20.8	48.9	96.0	64.1
	CNN	33.3	53.0	88.7	62.9
	LSTM	19.7	47.5	97.6	64.2
	Average	24.6	49.8	94.1	63.7

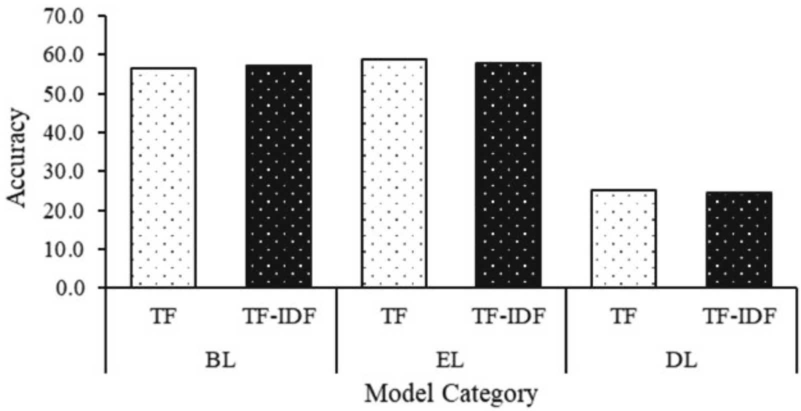


Fig. 8 Average Accuracy based analysis of COVID-19 related sentiments

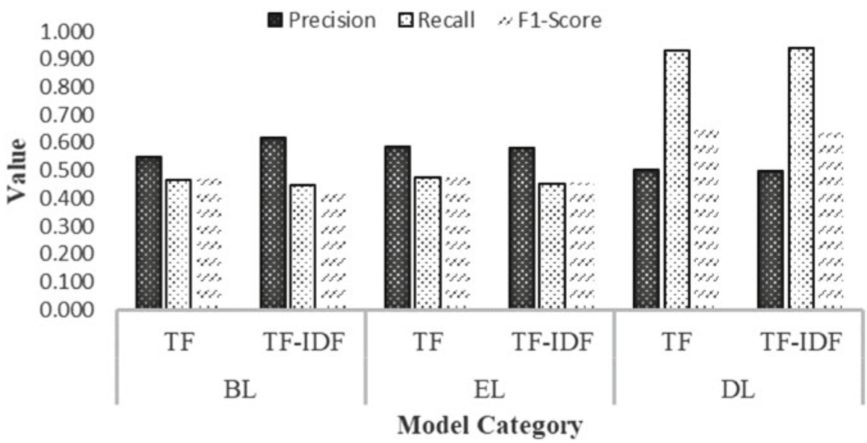


Fig. 9 Average Precision, Recall, F1-Score based analysis of COVID-19 related sentiments

4.2 Result and Analysis on Emotion Detection Task

4.2.1 Result and Analysis Using TF

From Table 7, it can be seen that the MNB, XGBoost and CNN are best performers in BL, EL and DL categories respectively. The best accuracy, precision, recall and F1-scores are respectively 37.5%, 29.4%, 24.5% and 22.6% (for MNB), 36.3%, 47.4%, 26.0% and 28.8% (for XGBoost) while 20.2%, 85.4%, 82.5% and 89.7% (for CNN).

4.2.2 Result and Analysis Using TF-IDF

From Table 8, it could be seen that SVM, with accuracy, precision, recall and F1-score of 35.1%, 37.3%, 21.1% and 19.2% respectively, accomplished better compared to BL algorithms. XGBoost, with accuracy, precision, recall and F1-score of 35.2%, 37.0%, 23.0% and 22.9% respectively, was best in EL category. Also, CNN was best in DL category. Accuracy (Fig. 10), precision, recall and F1-score (Fig. 11) for CNN was reported to be 18.7%, 92.2%, 89.4% and 87.8% respectively.

5 Conclusion

Average Precision, Recall, F1-Score based analysis of COVID-19 related emotions Social Media is platform for expressing your opinions, viewpoints, thought freely without any hesitation. During COVID-19 pandemic, world was physically disconnected due to COVID-19 restrictions but it is more connected in virtual environment.

Table 7 Results of algorithms using term frequency as feature weighing

Type	Model	Accuracy	Precision	Recall	F1-score
Baseline learners	LR	35.9	30.6	25.7	25.6
	KNN	24.9	15.5	18.7	15.9
	SVM	34.8	36.9	20.9	19.1
	MNB	37.5	29.4	24.5	22.6
	DT	27.3	20.8	19.3	19.4
	Average	32.1	26.6	21.8	20.5
Ensemble learners	RF	30.2	29.1	22.0	22.3
	XGB	36.3	47.4	26.0	28.8
	BG	31.6	36.4	20.0	26.9
	GB	32.3	34.8	24.3	26.7
	Average	32.6	36.9	23.1	26.2
Deep learners	ANN	11.6	83.7	99.7	90.4
	CNN	20.2	85.4	82.5	89.7
	LSTM	12.0	83.7	97.5	91.0
	Average	14.6	84.3	93.2	90.4

Table 8 Results of algorithms using TF-IDF as feature weighing

Type	Model	Accuracy	Precision	Recall	F1-score
Baseline learners	LR	37.1	35.6	23.3	22.3
	KNN	34.5	35.9	24.4	23.6
	SVM	35.1	37.3	21.1	19.2
	MNB	32.3	20.3	17.5	14.4
	DT	24.4	22.2	18.8	19.6
	Average	32.7	30.3	21.0	19.8
Ensemble learners	RF	29.3	22.7	19.1	17.9
	XGB	35.2	37.0	23.0	22.9
	BG	30.7	28.4	21.4	21.6
	GB	31.4	27.3	21.1	21.3
	Average	31.7	28.9	21.2	20.9
Deep learners	ANN	11.7	85.1	96.7	90.7
	CNN	18.7	92.2	89.4	87.8
	LSTM	11.1	82.7	98.4	90.5
	Average	13.8	86.7	94.8	89.7

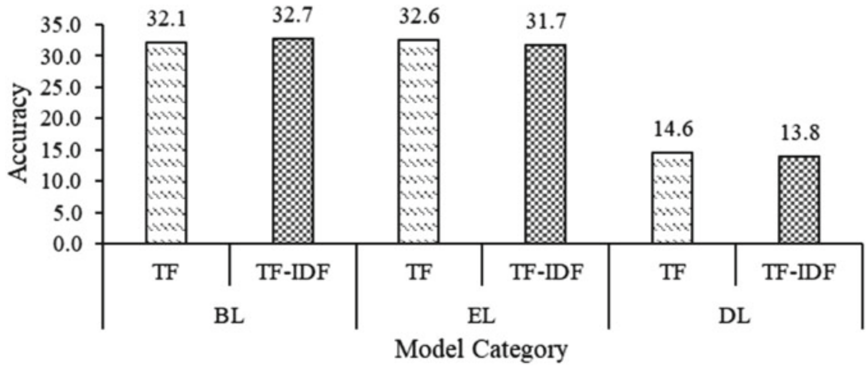


Fig. 10 Average Accuracy based analysis of COVID-19 related emotions

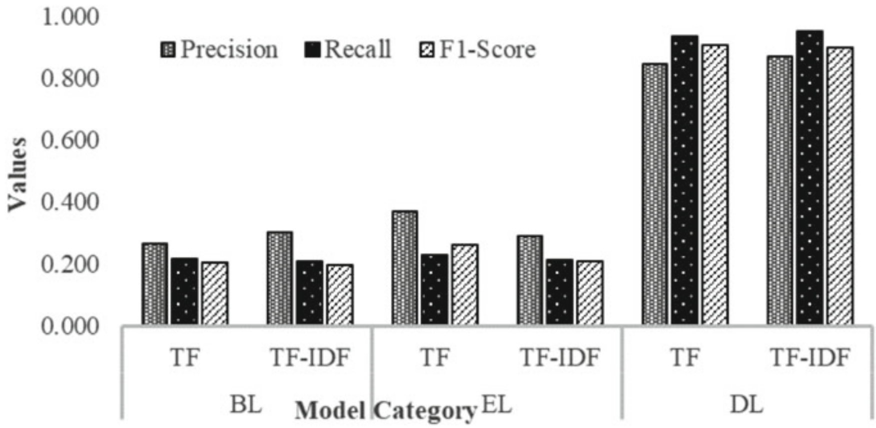


Fig. 11 Average Precision, Recall, F1-Score based analysis of COVID-19 related emotions

This research work was carried on corona virus outbreak using twitter data. The main focus of this study is to understand emotions and sentiments of people during COVID-19. This work helps to understand the people's perception about coronavirus and its impact on the public. The sentiments and emotions during the period were downloaded and the public's reaction towards the outbreak was analyzed. This dataset was passed through various pre-processing phases. Term frequency and term frequency-invers document frequency was used for feature extraction and feature weighing. To analyze sentiment and emotions, total 12 models were trained and tested using twitter dataset. These models were categorized into baseline, ensemble and deep learners. Results revealed that for sentiment analysis task, gradient boosting algorithm with term frequency as feature weighing (from ensemble learning models) outperformed all other models. Accuracy and Precision reported by gradient boosting model is 60.1% and 63.1%, respectively. For emotion detection task, Multinomial

Naïve Bayes model with term frequency performed better in comparison with other models.

References

1. Hu Y, Sun J, Dai Z, Deng H, Li X, Huang Q, Wu Y, Sun L, Xu Y (2020) Prevalence and severity of corona virus disease 2019 (COVID-19): a systematic review and meta-analysis. J Clin Virol 127:104371. https://doi.org/10.1016/j.jcv.2020.104371
2. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ (2020) Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 395:1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9
3. Adikari A, Nawaratne R, De Silva D, Ranasinghe S, Alahakoon O, Alahakoon D (2021) Emotions of COVID-19: content analysis of self-reported information using artificial intelligence. J Med Internet Res 23(4):e27341. https://doi.org/10.2196/27341
4. Machucal C, Cristian GC, Renato M, Toasa R (2020) Twitter sentiment analysis on coronavirus: machine learning approach. In: Proceedings of international symposium on automation, information and computing (ISAIC 2020), vol 1828
5. Dataset accessed from https: //www.kaggle.com/c/sentiment-analysis-of-covid-19-related-tweets/overview in April 2022
6. Behl S, Rao A, Aggarwal S, Chadha S, Pannu HS (2021) Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. Int J Disaster Risk Reduct 55:102101. https://doi.org/10.1016/j.ijdrr.2021.102101
7. Kwok SWH, Vadde SK, Wang G (2021) Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning analysis. J Med Internet Res 23(5):e26953. https://doi.org/10.2196/26953
8. Kabir M, Madria S (2021) EMOCOV: machine learning for emotion detection, analysis and visualization using COVID-19 tweets. Online Soc Netw Media 23:100135. https://doi.org/10.1016/j.osnem.2021.100135
9. Aygun I, Kaya B, Kaya M (2021) Aspect based Twitter sentiment analysis on vaccination and vaccine types in COVID-19 pandemic with deep learning. IEEE J Biomed Health Inf. https://doi.org/10.1109/JBHI.2021.3133103
10. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S (2021) An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 108:256–262. https://doi.org/10.1016/j.ijid.2021.05.059
11. Villavicencio C, Macrohon J, Inbaraj X, Jeng J, Hsieh J (2021) Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using naïve Bayes. Information 12(5):2021
12. Kazi NA, Shakib KM, Dhruba A, Khan M, Al-Amri JF, Masud M, Rawashdeh M (2021) Deep learning-based sentiment analysis of COVID-19 vaccination responses from Twitter data. Comput Math Methods Med 2021:4321131. https://doi.org/10.1155/2021/4321131
13. Yuming W, Stephen C, Erika P (2021) National leaders' usage of Twitter in response to COVID-19: a sentiment analysis. Front Commun 6. https://doi.org/10.3389/fcomm.2021.732399
14. Chandra R, Krishna A (2021) COVID-19 sentiment analysis via deep learning during the rise of novel cases. PLoS ONE 16(8):e0255615. https://doi.org/10.1371/journal.pone.0255615
15. Ridhwan K, Hargreaves C (2021) Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore. Int J Inf Manage Data Insights 1(2). https://doi.org/10.1016/j.jjimei.2021.100021
16. Marcec R, Likic R (2021) Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J. https://doi.org/10.1136/postgradmedj-2021-140685

17. Melton C, Olusanya O, Ammar N, Shaban-Nejad A (2021) Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: a call to action for strengthening vaccine confidence. J Infect Public Health 14(10):1505–1512. https://doi.org/10.1016/j.jiph.2021.08.010
18. Satu MS, Khan MI, Mahmud M, Uddin S, Summers MA, Quinn JMW, Moni MA (2021) TClustVID: a novel machine learning classification model to investigate topics and sentiment in COVID-19 tweets. Knowl Based Syst 226:107126. https://doi.org/10.1016/j.knosys.2021.107126
19. Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, Zhu T (2020) Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. J Med Internet Res 22(11):e20550. https://doi.org/10.2196/20550
20. Hazim M, Anuar NB, Ab Razak MF, Abdullah NA (2018) Detecting opinion spams through supervised boosting approach. PloS One 13(6), Article ID e0198884
21. Powers DM (2011) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. J Mach Learn Technol 2(1):37–63