Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
1st International Workshop on Mobile and Pervasive sensing for healthcare  
August 9-11, 2022, Ontario, Canada  

Human sentiments monitoring during COVID-19  
using AI-based modeling  
Areeba Umair\textsuperscript{a,}\textsuperscript{*}, Elio Masciari\textsuperscript{a,b}  
\textsuperscript{a}Department of Electrical Engineering and Information Technologies, University of Naples Federico II, Naples 80125, Italy  
\textsuperscript{b}Institute for High Performance Computing and Networking (ICAR), National Research Council, Naples, Italy  

Abstract  
The whole world is facing health challenges due to wide spread of COVID-19 pandemic. To control the spread of COVID-19, the development of its vaccine is the need of hour. Considering the importance of the vaccines, many industries have put their efforts in vaccine development. The higher immunity against the COVID can be achieved by high intake of the vaccines. Therefore, it is important to analysis the people’s behaviour and sentiments towards vaccines. Today is the era of social media, where people mostly share their emotions, experience, or opinions about any trending topic in the form of tweets, comments or posts. In this study, we have used the freely available COVID-19 vaccines dataset and analysed the people reactions on the vaccine campaign using artificial intelligence methods. We used TextBlob() function of python and found out the polarity of the tweets. We applied the BERT model and classify the tweets into negative and positive classes based on their polarity values. The classification results show that BERT has achieved maximum values of precision, recall and F score for both positive and negative sentiment classification.  

© 2022 The Authors. Published by Elsevier B.V.  
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)  
Peer-review under responsibility of the Conference Program Chairs.  

Keywords:  COVID-19, Vaccines campaign, Sentiments monitoring, Vaccine hesitancy, Social media data analysis, AI based modeling  

1. Introduction  
Coronavirus started spreading in Wuhan, China in December, 2019 [1],[11]. Statistics on Worldometer shows that COVID has effected the lives of 505,308,745 number of people, and has caused 6,225,378 number of deaths till 19 April 2022. Despite using the various preventive measure, the development of its vaccine was considered as permanent solution [3],[23]. Various pharmaceutical companies and other industries tried for developing the vaccines and many of them achieved their targets. Although, the vaccines were manufactures, the acceptance and willingness for vaccines was another hurdle in the prevention of COVID-19 [13],[6]. The research shows that the maximum control of COVID can be achieved my maximum supply of vaccines [20],[7]. The studies showed that most of the
population is not willing to accept the vaccines due to the misinformation spreading on social media [8], [10], [14]. Keeping in view, many governmental departments should extract and analyse people’s sentiments and reaction on vaccine campaign and ensure maximum uptake of vaccines [20], [21].

In recent years, data science is one of the growing fields due to the huge amount of data availability. The wide availability of data needs to transform in order to obtain useful information from it [5]. Data mining is a tool that treats and analyzes large volumes of data to discover relevant information for decision making that is usually invisible from a superficial level. The knowledge discovery process requires data mining to perform an iterative sequence of seven activities. Data mining is the activity in KDD, which analyzes specific patterns of any data [22].

With the paradigm shift towards social media, the availability of data is not a big deal. User-generated data is always available freely on different social media platforms, which can be acquired and use for various research purposes [2]. These tweets, comments and reviews, posted by users, can be analysed in order to highlight what people are thinking about any trending topic. Similarly, such data can be helpful in the days of pandemic and can help to investigate the people’s behaviour, feelings during the disease [12], [25]. Among all other social networks, twitter is the most famous due to its wide-spread use [4] and [19]. People make tweets for sharing their thoughts and using these tweets, one can easily investigate the people’s feedback [15]. The tweets are analysed using the methods of machine learning and deep learning [12]. One of these methods is sentimental analysis which uses polarity values of the given text [9][17].

In this research work, we have used tweets about vaccines and investigated the people’s feedback on vaccines campaigns. This research is helpful for agencies and policy making departments for designing the campaigns according to the sentiments of the people. The purpose of conducting the research was:

- Using freely available twitter data about COVID vaccines and categorize the text into different sentiment classes.
- Find the polarity of the tweets
- Using BERT model to achieve the classification of the sentiments of the people about vaccines.

The structure of this research article is as follows: The overall methodology of the proposed scheme is explained in section 2 while section 3 describes the discussion on the results. Finally, section ?? concludes our study and discuss the future work as well.

2. Methodology

Many methods of machine learning are available for text classification. The ML models use algorithms for the sentiment classification.

2.1. Proposed Scheme

The proposed scheme has three phases and is presented in Figure 1.

First phase is acquisition of tweets data and refining data by applying pre-processing steps. In this research, we used tweets related to COVID vaccines freely available on Kaggle website. First of all, we removed the hashtags, URLs and stop-word from the tweet text.

Second phase is finding the polarity values of the tweets. For performing the sentimental classification, the crucial thing is finding the sentimental polarity of the tweets. The polarity shows that how much the given text is neutral, positive and negative. We classified the tweets into seven classes on the basis of the polarity values. Hence, we used TextBlob() library of python for this purpose. The function returns the polarity values ranging from -1 to +1. Figure 2 explains the working principle of the TextBlob() function.

Third phase is sentimental (negative or positive) classification using BERT model. We encourage the use of deep learning for this purpose as it is inspired by the structure of human behavior in the form of artificial neural networks. There are many deep learning approaches proposed for different classification tasks. BERT, a transformer based machine learning model, is used in this research for sentimental classification. We have used it in our classification task because of its large number of application [16] [24].
2.1.1. BERT Architecture

Keeping in view the gaps in existing literature, we have proposed the use of BERT model for the sentiment classification of tweets. BERT uses unlabeled data for pre-training and based on the transformer. The Figure 3 shows the architecture of BERT model.
There are two special token found in BERT such as [SEP] and [CLS]. [SEP] is used for segment separation while [CLS] is used for classification. These tokens represent the first input of the classifiers. The probability of the classes can be found using the equation 1.

\[ P = \text{softmax}(CW^T) \]  
\[ (1) \]

2.1.2. Experiments

We tuned one separate model for positive sentiment classification and one separate model for negative classification. We used Recall (equation 3), Precision (equation 2) and F1 score (equation 4) matrices for model evaluation.

\[ \text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}} \]  
\[ (2) \]

\[ \text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}} \]  
\[ (3) \]

\[ F\text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \]  
\[ (4) \]

2.1.3. State-of-the-art models for comparisons

In order to evaluate our proposed approach, we have performed different experiments using state-of-the-art models. We have chosen the algorithms based on their frequently use in state-of-the-art [1], [12] and [18]. These algorithms are KNN, decision trees, SVM, random forest and Naive Bayes.

3. Results and discussion

3.1. Sentiment Polarity

The sentiment of the text can be identified using polarity values. In Table 1, polarity values and sentiment label of the sample tweets has been shown.

| Twitter data                  | Value of Polarity | Sentiment Category       |
|-------------------------------|-------------------|--------------------------|
| Headache after my first dose  | -0.25             | Weakly Negative          |
| Vaccines are the hope for better days | 0.675       | Strongly Positive        |
| Done with COVID-19            | 0                 | Neutral                  |

3.2. Vaccine Sentimental Classification

The Figure 4 and Figure 5 shows the results of positive sentiment classification and negative sentiment classification respectively.

Figure 4 and Figure 5 represent that our proposed model outperformed the other machine learning models for sentimental classification i.e. positive as well as negative classification. BERT resulted in 55 % precision, 69 % recall and 58 % F-score for positive classification. It resulted in 54 % precision, 85 % recall and 64 % F-score for negative classification. Out of all baseline neural network models, BERT outperformed all others. The reason for less accuracy of Random forest and decision tree is that they are famous for robust feature extraction and used for object recognition as well as image classification. Therefore, they showed less performance in our case. KNN is used to select neighboring labels with respect to sentiments.
4. Conclusion and Future Work

For the control of COVID-19, the development of vaccine is very necessary. Along with the preventive measures, vaccine helps to control the COVID on large scale. Meanwhile, the vaccine hesitancy is another big issue for COVID control and monitoring. We used TextBlob() to find the polarity of tweets and BERT model for sentiment classification of vaccines. The development of geographical information systems (GIS) approaches has enabled researchers to identify the spatial patterns in different domains i.e. education, health, poverty, earth observation etc. The GIS based approaches help to visualize the trends of people sentiments over map layers. The techniques of GIS such as hotspot analysis, buffering and link analysis help to identify the regions with higher value of positive and negative sentiments. Using such kind of methods and approaches in future can give more fruitful results.

References

[1] Adamu, H., Lutfi, S.L., Malim, N.H.A.H., Hassan, R., Di Vaio, A., Mohamed, A.S.A., 2021. Framing twitter public sentiment on Nigerian government COVID-19 palliatives distribution using machine learning. Sustain. 13. doi:10.3390/su13063497.
[2] Ceci, M., Corizzo, R., Fumarola, F., Ianni, M., Malerba, D., Maria, G., Masciari, E., Oliverio, M., Rashkovska, A., 2015. Big data techniques for supporting accurate predictions of energy production from renewable sources, in: Desai, B.C., Toyama, M. (Eds.), Proceedings of the
A performance comparison of supervised machine learning

Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., Choi, G.S., 2021. A performance comparison of supervised machine learning

Umair, A., Masciari, E., Habib Ullah, M.H., 2021. Sentimental analysis applications and approaches during covid-19: A survey, in: 25th

Umair, A., Masciari, E., 2022. Sentimental and spatial analysis of covid-19 vaccines tweets. Journal of Intelligent Information Systems, 1–21.

Sv, P., Ittamalla, R., Deepak, G., 2020. Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study .

Seale, H., Heywood, A.E., Leask, J., SHEel, M., Durrheim, D.N., Bolsewicz, K., Kaur, R., 2020. Examining Australian public perceptions and

Sv, P., Ittamalla, R., Deepak, G., 2020. Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study .

Sv, P., Ittamalla, R., Deepak, G., 2020. Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study .

Sv, P., Ittamalla, R., Deepak, G., 2020. Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study .

Sv, P., Ittamalla, R., Deepak, G., 2020. Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study .