Received: 31 March 2022 | Revised: 4 September 2022 | Accepted: 6 September 2022
DOI: 10.1002/cpe.7387

RESEARCH ARTICLE

WILEY

A novel TCNN–Bi-LSTM deep learning model for predicting sentiments of tweets about COVID-19 vaccines

Serpil Aslan

Software Engineering Department, Malatya Turgut Ozal University, Malatya, Turkey

Correspondence
Serpil Aslan, Malatya Turgut Ozal University, Software Engineering Department, Malatya, Turkey.
Email: serpil.aslan@ozal.edu.tr

Abstract

Many researchers in various disciplines have focused on extracting meaningful information from social media platforms in recent years. Identification of behaviors and emotions from user posts is examined under the heading of sentiment analysis (SA) studies using the natural language processing (NLP) techniques. In this study, a novel TCNN-Bi-LSTM model using the two-stage convolutional neural network (TCNN) and bidirectional long short-term memory (Bi-LSTM) architectures was proposed. While TCNN layers enable the extraction of strong local features, the output of these layers feeds the Bi-LSTM model that remembers forward-looking information and capture long-term dependencies. In this study, first, preprocessing steps were applied to the raw dataset. Thus, strong features were extracted from the obtained quality dataset using the FastText word embedding technique that pre-trained with location-based and sub-word information features. The experimental results of the proposed method are promising compared to the baseline deep learning and machine learning models. Also, experimental results show that while the FastText data embedding technique achieves the best performance compared to other word embedding techniques in all deep learning classification models, it has not had the same outstanding success in machine learning models. This study aims to investigate the sentiments of tweets about the COVID-19 vaccines and comments on these tweets among Twitter users by using the power of Twitter data. A new dataset collected from Twitter was constructed to be used in experimental results. This study will facilitate detecting inappropriate, incomplete, and erroneous information about vaccination. The results of this study will enable society to broaden its perspective on the administered vaccines. It can also assist the government and healthcare agencies in planning and implementing the vaccination's promotion on time to achieve the herd immunity provided by the vaccination.

KEYWORDS

Bi-LSTM, convolutional neural network, COVID-19 vaccines, deep learning, FastText, sentiment analysis

1 | INTRODUCTION

COVID-19 was first reported in December 2019 in Wuhan, China. It is an infectious disease caused by SARS-CoV-2. This disease, which was known as the Coronavirus Disease 2019 at the beginning, was declared a pandemic by the World Health Organization (WHO) on 11 March 2020.[1] COVID-19 is transmitted by small droplets formed when people cough, sneeze or speak. So, the first way to prevent this infectious disease is to maintain social distance from people. The COVID-19 pandemic has been a process of destruction, grief, and despair that resulted in the loss of millions of people's lives. According to the latest statistics, more than 428 million people have been affected by the coronavirus, and about 5.91 million deaths have occurred worldwide.[2] Countries worldwide have applied unprecedented and compelling pressures such as quarantine to slow the spread of the pandemic by keeping the number of cases under control during the pandemic and using available resources appropriately. Unfortunately, both the pressure applied and the rapid increase in the number of cases have caused the spread of negative emotions such as panic, depression, anxiety, fear, and stress in societies.[3] Although compelling pressures have been applied, the key to combating infectious diseases is undoubtedly the immunity acquired through having the disease in the past or vaccination. As in other epidemic diseases, immunity is gained against coronavirus bypassing the disease. However, because vaccines are optimally tailored to generate an effective immune response, the results of vaccine immunization are more predictable and therefore safer.

For the vaccine to have a protective effect, it must be administered by a large segment of the population. Over 115 vaccines are being researched and tested for the COVID-19 disease.[4] The immunity to be gained by vaccination may facilitate controlling the pandemic in conscious societies. People have suffered severe psychological injuries due to the intense anxiety, fear, and mixed emotions they experienced during the COVID-19 pandemic. Feelings became more complicated after the start of vaccinations. There are serious concerns and mistrust about vaccination in the community. Even with our family members, we often found ourselves in such discussions and conflicts of opinion. For instance, the rapid discovery of vaccines, insufficient research[5] on them, allergic problems caused by the vaccine, concerns about the necessity of taking the second dose, and so forth. Recently, it has been widely discussed in scientific articles. Society has been divided into three different groups on social media: pro-vaccine, anti-vaccine, and indecisive. The incomplete or incorrect comments made about vaccines on online social media platforms seriously impact individuals' vaccine attitudes. Inappropriate tweets shared about coronavirus vaccines reinforce the attitudes of individuals who have experienced vaccine rejection and cause their opinions to turn into vaccine rejection in individuals who have experienced ambivalence about vaccination. This is one of the biggest obstacles to gaining herd immunity. Studying the vaccine opposition and promoting the vaccine's trust by examining the sentiments and opinions about the vaccine can encourage vaccination among the population.

In this context, Twitter is the most popular social media resource where people share their feelings and opinions on the subject.[6] However, accessing individuals' ideas about vaccines is limited to our immediate focus as online users. It is almost impossible for a person to read and digest all tweets about the COVID-19 vaccines. Rumors about any topic discussed on Twitter can be easily categorized according to two AI techniques algorithms: machine learning, deep learning. There are essential differences between these two methods. While machine-learning algorithms require a feature selection process, deep learning algorithms can extract feature selections directly from the textual content without human assistance.[6] Deep learning algorithms consist of several interconnected layers. The input given to each layer is interpreted and transferred to the next layer. Using natural language processing (NLP) techniques, sentiment analysis, and word cloud visualization techniques will facilitate the study of complex tweets of enormous size. However, some aspects of the existing methods still need improvement.

Numerous publications have analyzed the public's emotions and opinions over the COVID-19 pandemic and the COVID-19 vaccinations on Twitter. Mahdikhani[7] performed Twitter sentiment analysis for different stages of the COVID-19 pandemic from the first times of the pandemic process to the last. They investigated public emotions and opinions during the various stages of the COVID-19 pandemic using deep learning-based sentiment analysis techniques. The conclusions have led to the argument that because popular tweets are more widely read and discussed, they should be the first to be found and studied. Okango et al.[8] analyzed the effect of the COVID-19 pandemic process on society with dictionary-based sentiment analysis techniques. They revealed that the strict rules of the pandemic process and material shortages in the market create serious mental health problems in society. Valle-Cruz et al.[9] investigated the effect of sentiment poles of tweets shared on Twitter social media platforms on world financial indices during the pandemic process. They used tweets about COVID-19 and H1N1 pandemics in this study. Chinnasamy et al.[10] used machine learning-based approaches to analyze the tweets on Twitter to understand how the public feels about the COVID-19 vaccines. The results demonstrated that the population's emotional response to vaccine use was generally either neutral or favorable. In addition, experiments showed that initial fear and anxiety toward vaccines decreased somewhat over time.

This study investigates the sensitivity of tweets and comments on these tweets about COVID-19 vaccines via deep learning-based NLP techniques by using the power of Twitter data. This paper proposes a novel hybrid model (TCNN–Bi-LSTM) based on CNN and Bi-LSTM models to detect rumors about the COVID-19 vaccines on Twitter. The proposed model connects a two-stage convolutional CNN (TCNN) layer to Bi-LSTM. This study trained the proposed TCNN-Bi-LSTM model via the FastText word Embedding technique. Moreover, to show the success of FastText word embedding, all compared models were trained and tested with TF-IDF and Glove techniques. The k-fold cross-validation method was applied with $k = 10$ throughout the training and testing of the classifiers to enhance their accuracy. When the experimental results are examined, it is seen that the proposed method has higher performance compared to other models.

1.1 Motivation

Today, social media has become the most preferred tool for spreading false information and allowing the rapid spread of correct information. Furthermore, with the widespread use of the Internet, millions of people share their feelings and thoughts using social media tools daily. Therefore, social media tools have become an enormous resource for obtaining user opinions today. For this purpose, Twitter has become an essential social media tool where general users share their thoughts.[11] On Twitter, users share tweets that express their feelings and opinions, using a maximum of 140 characters.

During the quarantine period implemented by governments these days as we fight against the coronavirus, social media platforms have been inevitable for societies that have to spend time at home. In such a challenging period, individuals' social platforms to express themselves should guide us properly. However, unfortunately, contrary to expectations, many individuals sharing inappropriate content and false information through social media platforms is substantial.

Since pandemics such as coronavirus occur once in a lifetime, the methods of combating such pandemics have not been fully determined. While some countries are successful in the fight against the pandemic, others have suffered social and economic problems that would take time to solve. Some society members rejected vaccination during the vaccination process, which is considered the key to the fight against the pandemic. Inappropriate tweets about vaccination reinforce the attitudes of individuals who reject vaccination and cause individuals who are indecisive about vaccination to reject vaccination.

This study investigates the sensitivity of tweets about COVID-19 vaccines among Twitter users and the comments made on these tweets. Also, it detects inappropriate, incomplete, and incorrect posts by using deep learning-based NLP techniques. After detecting negative sentiments on social media, tagging posts containing false information with content warnings can help reduce the impact of incorrect information. The results obtained from this study will enable society to broaden its perspective on the administered vaccines. Furthermore, this study will guide the government and health institutions to plan and implement the vaccine's introduction on time to achieve the herd immunity provided by the vaccination.

The contributions of this study are summarized as follows:

- Social media users often use typographical errors (typos) and abbreviations in tweets. A typo can completely change the point of view of the thought to be conveyed. Therefore, powerful preprocessing tasks were applied to clean the data, remove noise, and lemmatize.
- To extract high-level features used in sentiment analysis in the study, the FastText word Embedding technique is preferred. Due to FastText being pre-trained with location-based and sub-word information features, it can process hidden features in language and out-of-vocabulary words. This model also creates meaningful word representations by retaining the word order information of tweets. Experimental results prove the successful performance of the utilized word embedding technique.
- Although strong local features can be extracted from CNN layers thanks to the convolution and pooling layer, correlation sequences cannot provide detailed learning. In addition, while Bi-LSTM architecture cannot extract local features successfully compared to CNN, it is successful in learning correlation sequences. Therefore, using CNN alone or Bi-LSTM alone may be a weak model while achieving optimal classification results. Thus, a hybrid CNN and Bi-LSTM model named TCNN–Bi-LSTM was proposed in this study. A two-stage convolutional layer (TCNN) was used to extract more meaningful local features by using the CNN model. In the TCNN network model, the output of the convolutional layer feeds the next convolutional layer as input. So that stronger local features are extracted. Later, the output of TCNN feeds the Bi-LSTM model, which is known for its success in capturing long-distance dependencies. Finally, these two architecture were combined in a single hybrid TCNN-Bi-LSTM model.
- The two-stage TCNN model captures and learns high-level local information sufficiently in the proposed new model. In contrast, the Bi-LSTM model remembers forward-looking information and learns extensive correlation sequences.

The remaining of this study is organized as follows. First, the related works on sentiment analysis are introduced in Section 2. Then, dataset collection, Preprocessing steps, and the architecture of the proposed model are described in Section 3. Next, the experimental results are provided in Section 4. Finally, the results of the study are discussed in Section 5.

2 | RELATED WORKS

Recently, sentiment analysis (SA) has become one of the most popular study topics in NLP. Subjective information intended to be transmitted over a text via SA is automatically extracted.[12] Thanks to sentiment analysis, sentiment extraction, sentiment classification, summarization, and so forth. many operations are performed. Sentiment analysis started as a document-level classification task,[13,14] and has been used at the sentence level[15,16] and more recently at the abstract-based level.[17,18] Common sentiment analysis models rely heavily on sentiment dictionaries. A sentiment dictionary is a group of words in which each word is labeled according to its positive and negative subjective orientation. In addition to all this, different types of dictionaries can be classified in two ways: Semantic orientation labeling (in which words are grouped into having positive or negative meaning) or quantitative scoring according to predefined rules. LIWC[19] and HULIUO[20] are dictionaries of commonly used context-free words. On the other hand, ANEW,[21] SenticNet,[22] and SentiWordNet[23] are dictionaries with quantitative scoring based on sensitivity density. There are many challenges due to the sentiment dictionary. Because of the several difficulties of human languages, such as endless grammatical variations, idioms, slang, or misspellings, automatic analysis of natural language becomes difficult. These reasons have led to the development of machine learning-based approaches to perform sentiment analysis. Texts are often converted into features using a bag of words in machine learning-based approaches.[24] Then, these attributes, which are generated using complex machine learning techniques, are fed with classifiers such as Naive Bayes (NB), Decision Trees (DT), or Support Machines (SVM).[25] Machine learning-based approaches were later replaced by deep learning-based approaches. Deep learning-based approaches provided a significant performance increase compared to machine learning-based approaches in experiment results of studies carried out in this area.[26] Many approaches have been proposed in this field. Klachbrenner et al.[27] proposed the convolutional neural network model to analyze input sentences with variable lengths. Irsoy et al.[28] proposed an recurrent neural network model (RNN), which consists of interrelated inputs and is capable of remembering, for sentiment classification. Taio et al.[29] improved the RNN architecture and proposed the LSTM model with feedback links. Schuster and Paliwal[30] improved the traditional LSTM networks and proposed the Bi-LSTM model that uses two different LSTM networks. In this study, sentiment analysis was performed based on sentence-level sentiment analysis. In addition, a hybrid model based on a two-stage TCNN-BiLSTM was proposed as a classifier.

In recent years, Twitter have been frequently used for the dissemination of both authorized information and false information. Applying sentiment analysis on Twitter has been a trend platform for researchers to conduct scientific experiments and their potential applications.[31] Analyzing tweets on these platforms and performing sentiment analysis will provide a great convenience in detecting tweets that will arouse negative feelings in society. Joshi et al.[32] made comparisons using SVM, NB, and Maximum Entropy methods to analyze movie reviews on Twitter. Parikh et al.[33] applied two Unigram NB models, the Bigram NB model, and the Maximum Entropy model to classify tweets. Go et al.[34] performed sentiment analysis using NB, Maximum Entropy, and SVM machine learning classifiers with different data representation architectures such as unigram, bigram, and n-gram. In their study, they reported that SVMs provided better results compared to other classifiers. According to the studies in the literature, Naive Bayes and neural networks are commonly used to classify text in datasets with small text sizes, while in datasets with increasing data sizes, deep learning architecture is mostly used to increase accuracy and to process the neural network created with big data. Through the deep learning approach, the sentimental features in the text are detected by automatic definitions. This study proposes a model that conducts sentiment analysis predictions on a big dataset collected from Twitter using deep learning architecture.

Word embedding in another saying word representation plays a vital role in generating continuous word vectors based on the word's meaning in the dataset. Word embedding techniques that capture the semantic and syntactic features of words can also measure word similarity. Measured word similarities are commonly used criteria in NLP operations. There are various word embedding techniques in the literature. A few of them: Bag of Words,[35] Word2Vec,[36] Glove,[37] TF-IDF,[38] and LDA.[39] Fu et al.[40] used Word2Vec to represent datasets as input to the recursive autoencoders. Abid et al.[41] proposed the Glove-Bi-GRU-CNN model, which combines the Glove word embedding technique with RNN and CNN architectures to strengthen the feature set and improve its performance in sentiment analysis. Word embedding techniques play a vital role in emotion classification. However, most of the word embedding techniques in the literature can cause a lack of semantic information and ignore the emotional information in the text.[42] In 2013, Mikolov et al. FastText, recommended by,[36] is one of the most popular word embedding approaches in recent times. The proposed model is a Word2Vec based model. With this method, the pipeline of Word2Vec is changed to letter-based n-grams. Zhao and Mao[43] proposed a novel model, which combined TF-IDF and Word2Vec word embedding approaches. The experimental results of the study prove the success of the model. The proposed method, which was a combination of feature weighting and word vector, proved to perform the tasks of extracting meaningful words. Kim and Yoon[44] proposed a new sentiment analysis model using CNN and Bi-LSTM architectures. In the proposed model, they first provided the sentiment dictionary and word embedding dictionaries created by Word2Vec. Then, they constructed the CNN-Bi-LSTM model trained Word2vec technique. Al-Sarem et al.[45] proposed a model that took the same word embedding vector technique as input to three parallel CNN models, then combined them and fed the LSTM model with the result. The researchers tested each of the Word2Vec, Glove, and FastText word embedding techniques separately in their proposed model. Kamyab et al.[46] proposed a new approach word embedding approach to extract a strong feature set in sentiment analysis. The proposed approach based on TF-IDF and Glove. The reviewed studies prove that the proposed model of using the correct word embedding technique can improve text classification performance.

Social media platforms and media tools should provide constructive and impartial information with proven accuracy that will arouse positive emotions rather than presenting unfounded information that will divide the society and arouse negative feelings such as fear, anxiety, and stress in the society. Many studies were conducted on the analysis of such shared information. COVID-19 is a great affliction for the whole world. Since the news about COVID-19 concerns all humanity, it has become more of the agenda than all other events. In addition, this news may contain fake, unverified information and information that may be the subject of people's prejudices. Therefore, studies on determining the negativities in the news about COVID-19 are timely and appropriate. Chakraborty et al.[3] proposed a novel model to analyze tweets about the coronavirus pandemic on the Twitter social media platform. In the proposed study, they argued that the tweets with negative content created emotions such as fear, horror, and sadness that have great psychological effects on individuals, and that detecting these tweets would provide great advantages in the fight against the pandemic. Barkur et al.[47] analyzed a dataset consisting of 2400 tweets they collected regarding the COVID-19 news in India. Christopher et al.[48] investigated what kind of effects the COVID-19 pandemic process had on human psychology. They concluded that due to the COVID-19 news, people became aggressive, tense, and their depression levels increased, as a result of the study. Sural et al.[49] revealed that intelligence, which they defined as continuous emotional intelligence, was related to the problematic use of social media and was directly related to directing users to current issues on social media. This study argued that users with continuous emotional intelligence scores use social media platforms as a coping strategy to overcome the problems they experience in their private lives. Xue et al.[50] collected tweets about COVID-19 using 25 different tags between 1

March 2020 and 21 April 2020. The study was carried out using a dataset consisting of 4 million tweets. Popular topics in tweets were identified using the latent dirichlet allocation (LDA) algorithm. Feng et al.[51] used the BERT model, which focuses on single and multiple tags, to examine the emotions in tweets shared on Twitter. The proposed model focuses on the use of emojis for users to express their emotions. The study of Hornung et al.[52] argued that there was a proportional relationship between emotional intelligence and the use of Facebook. In the proposed study, a dataset consisting of different age groups was collected. According to the experimental results applied on this collected dataset, it was reported that there was a positive correlation between emotional intelligence and Facebook use for the young group in the dataset. However, there was a negative correlation for the elderly group at the same time. Depoux et al.[53] discussed several measures to be taken regarding the social media panic that emerged in the fight against the coronavirus pandemic. With the increase in the use of social media during the stressful quarantine days that many societies were exposed to, the number of posts that would create a panic effect has increased on social media. It was emphasized that such posts should be prevented by the measures to be taken by the Ministry of Health and its affiliates, and using social media in a wise manner would help individuals keep their psychological balance. Jelodar et al.[54] used NLP for modeling the topic related to users' posts about COVID-19 on social media. Also, they proposed a new model based on LSTM and RNN architecture for classification. Pano et al.[55] proposed a VADER-based model. They aim to analyze the impact of the COVID-19 pandemic process on Bitcoin prices.

Vaccination is undoubtedly one of the most important inventions in the fight against epidemic diseases. Negative news about the vaccines and anti-vaccination social media posts can reduce the rate of vaccination and prevent herd immunity, which is based on vaccination. For example, posts containing false and incomplete information about the negative effects of vaccines are often manipulated by anti-vaccine to fuel their actions.[56] Conspiracy theories were produced against vaccine development even before the development of COVID-19 vaccines started. Therefore, the sentiment of posts about the COVID-19 vaccination process shared on social media platforms should be examined and reviewed based on factors influencing vaccine acceptance to promote vaccination. Praveen et al.[57] examined their feelings and thoughts about the COVID-19 vaccines by using Twitter shares of individuals living in India. Shamrat et al.[58] proposed a sentiment analysis model based on the KNN machine-learning algorithm to analyze each COVID-19 vaccine sentiment on Twitter. Alam et al.[59] proposed a novel model based on LSTM and Bi-LSTM deep learning architectures to analyze people's attitudes and behaviors toward coronavirus after the beginning of vaccination. Yousefinaghani et al.[60] conducted a statistical analysis to determine the public's feelings and thoughts about COVID-19 vaccines.

3 | METHOD

This section explains the steps of the dataset collection and preprocessing techniques, the sensitivity analysis, and the proposed TCNN-BiLSTM model for classification. The raw dataset was first collected in the study, and then the preprocessing process was applied to the collected raw dataset. The data collection and preprocessing steps followed in the study are presented in Figure 1. Finally, the proposed model was performed on the preprocessed dataset.

3.1 | Dataset collection

Twitter, the world's largest social media platform with 199 million daily active users as of the first quarter of 2021.[61] The data were collected from the tweets shared publicly on Twitter. Public tweets about COVID-19 vaccines shared in English for this dataset were collected. In the study, MAXQDA,[62] a qualitative data analysis tool, was used to collect Twitter data. On this platform, tweets, retweets, and comments on the COVID-19 vaccine posted between 16 June 2021 and 1 November 2021 were collected. First, the most tweeted hashtags about the vaccine on Twitter were determined. A combination of the "*#covid19aşı*, *#biontech*, *#Pfizer/BioNTech*, *#Moderna*, *#Sinovac*, *#covid19vaccine*, *#CovidVaccine*, *#coronavac*, *ve #vaccinesaveslives*, *#vaccinevictims*, *#getvaccinated*, *# vaccinesideeffects*" search terms was used to reach the target tweets. Finally, the collected data were converted to the CSV format to use in pre-processing and emotion classification steps. The Boolean operators "AND" and "OR" indicated that tweets containing the words with the root of "*coronavirus*" or "*COVID*" could also be searched, as well as words with the root of "*vaccine*". Finally, a dataset of 368,001 tweets in English was collected. However, two datasets, the public sentiment dataset of COVID-19 vaccinations, which were anonymously published at Kaggle, were joined with the original dataset to enhance the dataset and improve analysis. The total number of tweets in the dataset was 973,613 after adding new tweets. Then, the pre-processing stage, the details of which are given in the next section, was carried out to ensure the accuracy of this dataset.

3.2 | Data cleaning and pre-processing

Data collected from the Twitter social media platform are often not clean. *Data cleaning* is a text mining process applied to clean the components in the text that are difficult to understand or analyze. They may include many unnecessary special characters, expressions, links, tags, emojis, and so

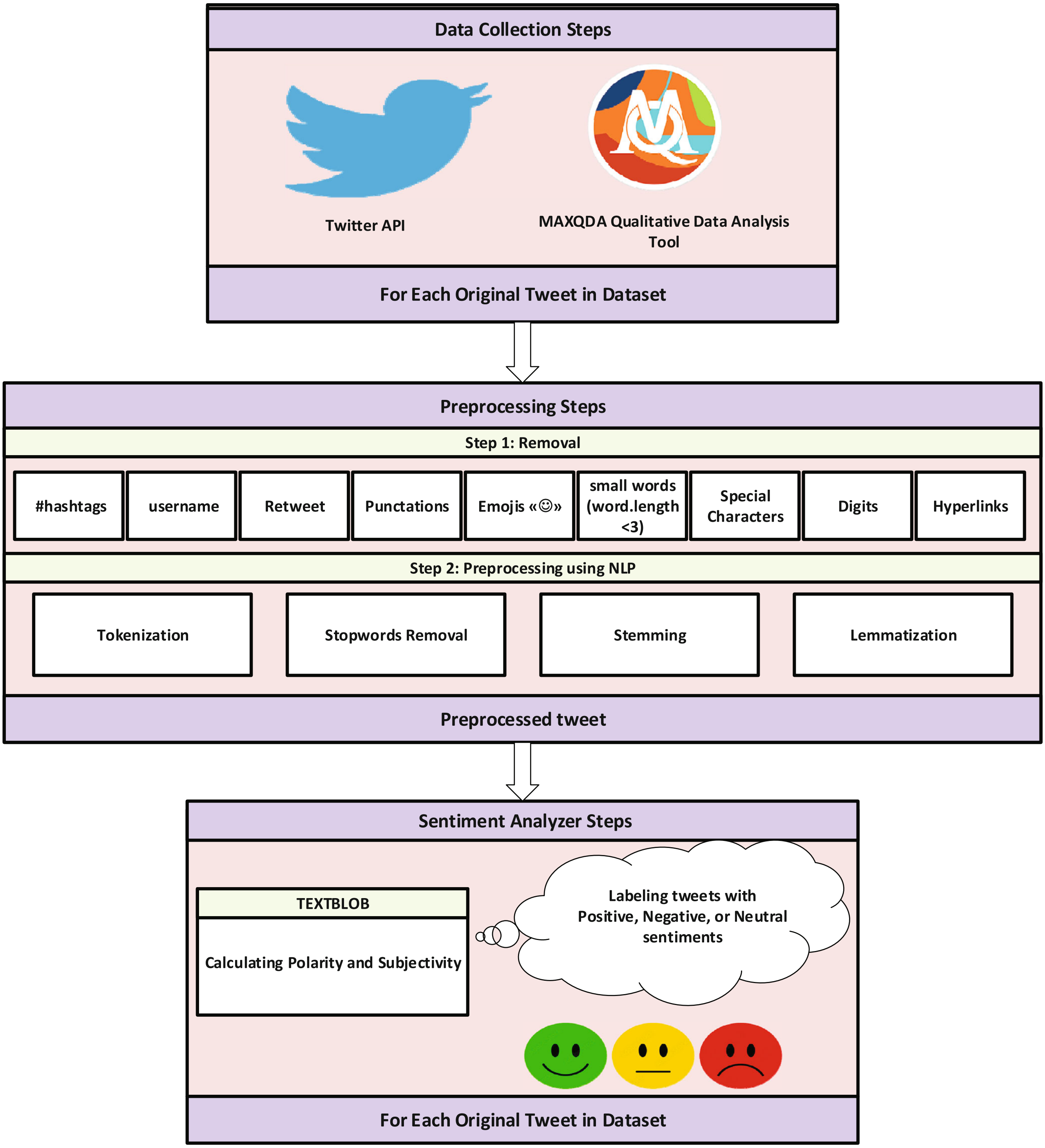
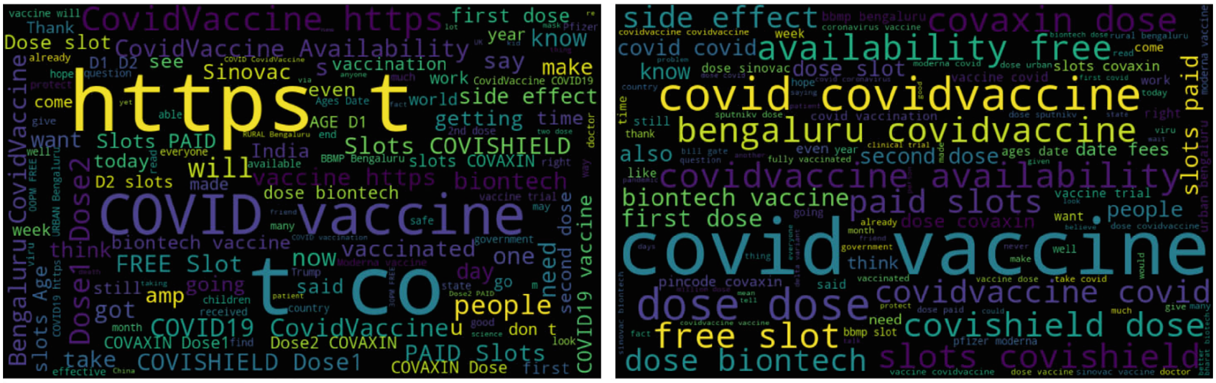


FIGURE 1 Dataset collection and preprocessing steps



(A) (B)

FIGURE 2 (A) Word cloud of the raw tweets, (B) word cloud of the preprocessed tweets

forth, that may adversely affect the experimental results during the analysis stage. Such characters do not contain meaningful information about the text. Even it causes difficulties in the emotion analysis phase. At this stage, the dataset was preprocessed by applying the operations given in Figure 1.

Figure 2 shows a word cloud of the collected tweets before pre-processing steps. The word cloud is one of the most popular data visualization techniques used to represent textual data. Essential and common textual data can be highlighted using the word cloud notation. Pre-processing steps aim to reduce the number of words in the text without changing the basic meaning of the text. As seen in Figure 2, the raw dataset includes unnecessary words and expressions that will not be used in the sentiment analysis stage. Therefore, the presented word cloud representation validates that pre-processing must be performed before analyzing any data. As seen in Figure 2, the word space has shrunk considerably after cleaning up unnecessary and meaningless words.

3.3 | Polarity and subjectivity

TextBlob is a library that manipulates natural language processing (NLP) data. TextBlob executes tasks via the natural language ToolKit (NLTK). NLTK provides a simple API for lexical NLP tasks such as sentiment analysis, root extraction, classification, and translation.[63] TextBlob is a rule-based library. It uses a special dictionary that defines positive and negative words (*sloria/TextBlob on GitHub at en/ en-sentiment.xml*[64]). TextBlob is called after converting the data to the appropriate format for performing sentiment analysis. TextBlob library returns two properties, “Polarity” and “Subjectivity”.[64]

3.3.1 | Sentence subjectivity

The speaker’s thoughts, feelings, or likes are subjective sentences. Subjective judgments cannot be proven because their truth or falsehood varies from person to person. Subjective sentences usually reflect personal feelings such as beliefs, desires, opinions, doubts, joys, fears, and so forth. On the other hand, the objective sentence is the sentence that contains a definite diagnosis that is accepted by everyone and does not vary from person to person. Subjectivity is a variable parameter that takes a value in the range [0, 1]. The closer the subjectivity value of the sentence is to zero, the more relevant the sentence is to the facts. As the value of subjectivity increases, it gets closer to being an opinion. On the other hand, the objective sentence is the sentence that contains a definite diagnosis that is accepted by everyone and does not vary from person to person.

3.3.2 | Sentence polarity

Polarity refers to written or spoken language’s positive, negative, and neutral emotional orientations. The polarity takes a value in the range $[-1, 1]$. Polarity scores of -1, 0, and 1 represent negative, neutral, and positive expressions, respectively.

To explain how polarity and subjectivity values are calculated:

In the TextBlob, the text given to it as input is first represented as a word bag (BoW). Then, after assigning individual scores for each word, the overall sentiment is calculated by averaging these scores. In addition, words that indicate negation in the input data also reverse the polarity of the sentence. Each word in the dictionary used by the TextBlob library has a particular score. These:

TABLE 1 An example of polarity, subjectivity, and intensity[65]

Word	Polarity	Subjectivity	Intensity
Great	1.0	1.0	1.0
Great	1.0	1.0	1.0
Great	0.4	0.2	1.0
Great	0.8	0.8	1.0

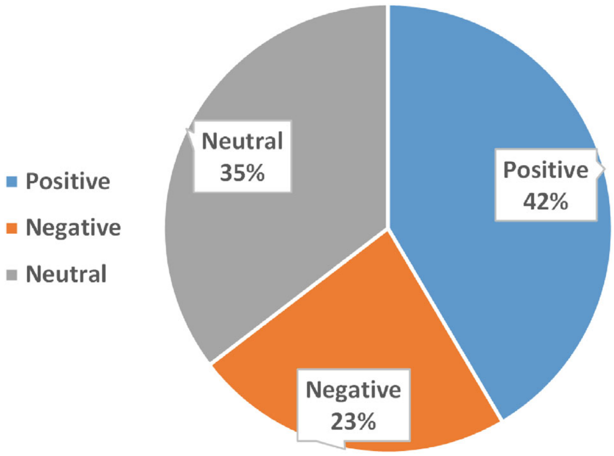


FIGURE 3 Sentiment distribution of COVID-19 vaccines

$$\text{Polarity : positive or negative } (-1.0 \rightarrow +1.0)$$
$$\text{Subjectivity : objective or subjective } (+0.0 \rightarrow +1.0)$$
$$\text{Intensity : modifies next word?}(\text{x}0.5 \rightarrow \text{x}2.0)$$

Table 1,[65] it shows an example of a word found in the special dictionary that TextBlob uses to calculate individual scores for each word in the word bag: polarity, subjectivity and intensity.

As seen in Table 1, the word “great” in the special dictionary used has polarity, subjectivity, and intensity values for four different definitions. When the word “great” is given as input for the TextBlob tool, it is calculated as polarity = 0.8 and subjectivity = 0.75. Because TextBlob uses the “averaging” technique applied to the polarity scores while calculating the polarity score for each word. The same process applies to each word in the BoW obtained from the input data. After the individual word scores are calculated, the average of the polarity scores of each word in the BoW is calculated at once to calculate the sentence polarity.

Of the dataset, including 973,613 tweets, the remaining 955,526 tweets after preprocessing stage were analyzed using TextBlob and divided into three classes: Positive, negative, and neutral. Figure 3 shows the number of tweets separated by the classification tags. According to the results of the analysis, 396,434 (41.49%) tweets were found to be positive, 220,921 (23,12%) tweets were found to be negative, and 338,171 (35,39%) tweets were found to be neutral. While positive tweets, which comprise the majority in sentiment distributions, show pro-vaccine groups, neutral tweets show confusion and uncertainty about COVID-19 vaccination procedures. The number of negative tweets about the anti-vaccine trend is substantial. Negative tweets about coronavirus vaccines may reinforce the attitudes of individuals who have experienced vaccine rejection and cause their opinions to turn into vaccine rejection in individuals who have experienced ambivalence about vaccination. These tweets will undoubtedly harm vaccination works. In this paper, we separated tweets into two classes label_1 (represent positive tweets denoted by 1) and label_2 (represent negative tweets denoted by 0).

In the proposed method, certain words or terms in tweets were divided into polarity groups by a word cloud. Figure 4 shows the terms in tweets tagged with positive, negative, and neutral sentiments. As can be seen in Figure 4, which represents the emotions that Covid-19 vaccines have left on society; (A) the word cloud of positive tweets represents words that contain positive emotions felt by pro-vaccine individuals regarding vaccination and the Covid-19 pandemic process, (B) The word cloud of negative tweets consists of words that contain negative emotions shared by anti-vaccine individuals, (C) the word cloud of negative tweets consists of words that do not contain any positive or negative emotions shared by individuals with vaccine indecision. Since “COVID, dose, vaccine, vaccinated, COVID vaccine, vaccination” popular terms that

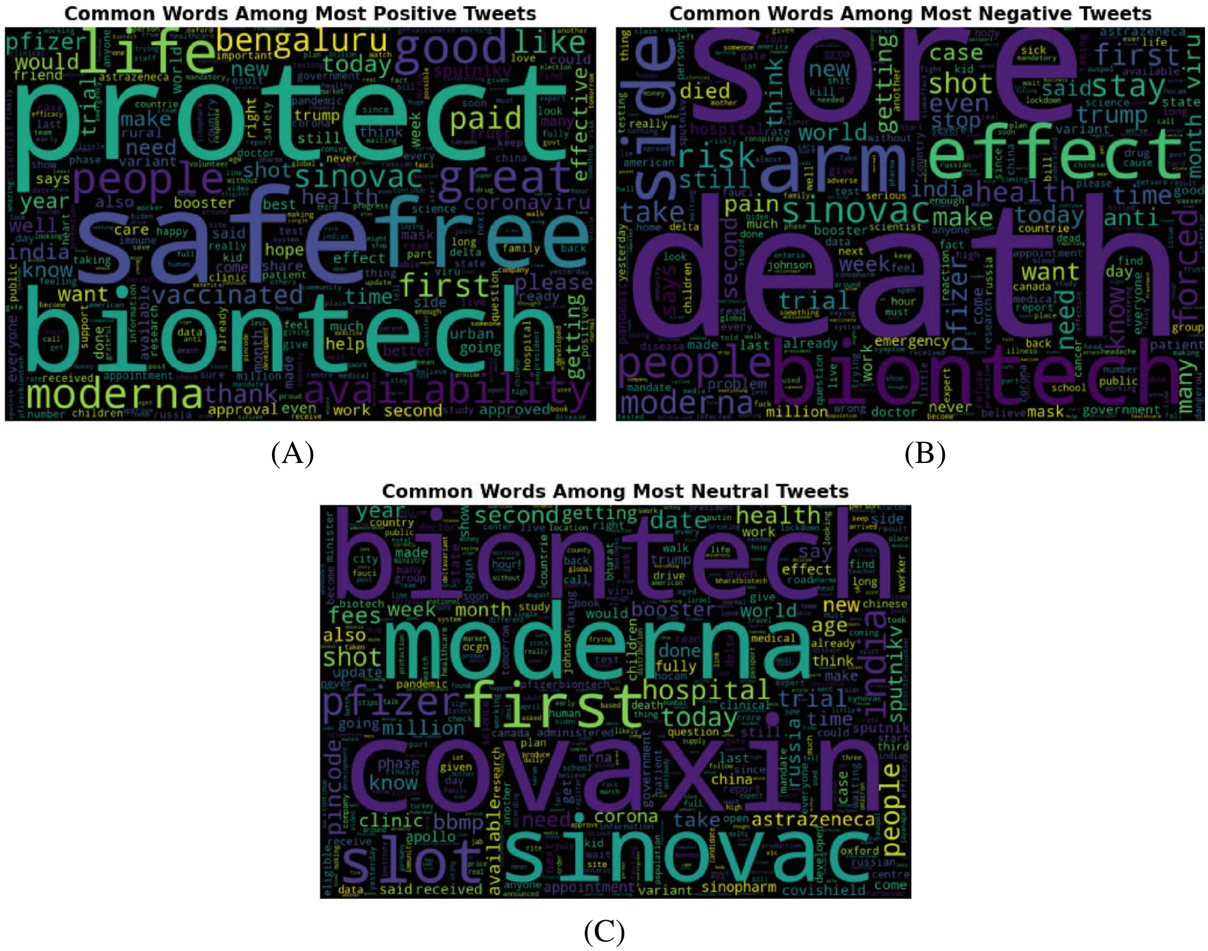


FIGURE 4 Word cloud of the common words among most (A) positive tweets, (B) negative tweets, and (C) neutral tweets

have the highest frequency and they are not necessary highlighting any polarity from word clouds, they cleaned from the polarity of tweets word clouds.

3.4 | Word embedding representation

Word embedding representations are an effective method used in text classification applications with deep learning and machine learning architectures. Thanks to word embedding representations, text documents are more clearly represented. The proposed model was trained with TF-IDF, Glove, and FastText word embedding representation schema in this study.

TF-IDF[46] identifies words that occur frequently in a given text and words that are not common in the remaining dataset. In this way, the texts can be evaluated as a short document. The multiplication of TF and IDF provides a measure of how often the word is included in a document multiplied by how unique the word is, and this is the TF-IDF measure. Words such as "this", "what" or "if", which are common in every text, do not make much sense for these textual sentences, even though these words appear many times, they are in the lower ranks in the order of density.

Global vectors for word representation (GloVe) was suggested by Pennington et al.[37] as an alternative to the Word2Vec approach to obtain word vector representations. It is an unsupervised scheme used to obtain vectorial representations of words. It utilizes the statistics of words used together for generating the word representations. A pre-trained Glove model built using approximately 2 billion tweets was used to generate a 200-dimensional word vector matrix.[46]

FastText is a Word2Vec-based model developed by Facebook for text classification in 2016.[36] It is a computationally efficient scheme for generating word embedding representation vectors from text documents. With this model, texts or words are converted into continuous vectors. The difference of this method from Word2Vec is that it breaks down words into several letter-based "n-grams" instead of giving them as individual inputs to the artificial neural network. Thus, the proximity of meaning that cannot be captured with Word2Vec can be captured with this method. As with

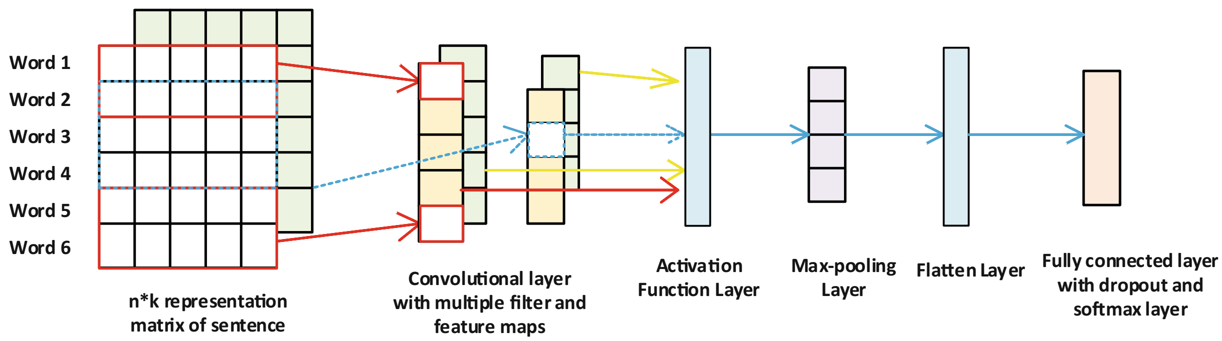


FIGURE 5 Basic convolutional neural network (CNN) architecture for text classification

Word2Vec, it offers two models in FastText: Skip-gram and CBOW. Skip-gram uses closed neighbor words to guess a target word, while CBOW uses a perspective that covers all the words in the content to guess the target word.[36] Both methods create a text file with a numerical representation of the learned words (vectors). This study used a pre-trained FastText word embedding technique (fastText Wiki + news)[66] which constructed by CBOW architecture was used.

3.5 | CNN

CNN, also called covnets, are deep learning architecture types with standard parameters. In recent years, they have been the focus of attention of many researchers. A covnet consists of a series of interconnected layers, where each layer can transform one volume into another through a differentiable function.[67] CNNs are widely used in NLP, image recognition, and other fields to recognize local features with a convolutional kernel.[68] Also, CNN architectures can learn features that will perform the classification automatically.

The unsupervised CNN architecture is a multilayer feed-forward neural network. CNN consists of various layers. In the input layer, the raw input of the data (can be image, text, etc.) is held. The convolutional layer is the most fundamental building block of CNN architecture. It can be applied as a single or multiple. This layer is the layer where the properties of the input data are detected. The activation function layer is the layer where activation functions (Softmax, Sigmoid, ReLU, tanh, etc.) are applied to the outputs of the input layer or the convolution layer. Activation functions play a significant role in neural networks. Since these can directly affect the model's performance, the pros and cons should be investigated in detail. The pooling layer, which can be max pooling or average pooling, is the layer applied to reduce the volume size, increase the computational efficiency, and reduce the computational complexity. The flattened layer feeds the output of the pooling layer. It also maps it to the next layer. The fully connected layer is the last and most important layer of CNN. In this layer, class scores are calculated. Figure 5 shows the basic CNN architecture used in text classification studies.

As seen in Figure 5, an n-dimensional sentence S is represented by the matrix $x = [w_1, w_2, \ldots, w_n]$. The i^{th} element in w_i is the k-dimensional vector representation of the i^{th} word in sentence S. The matrix $x \in R^{n \times d}$ is the input to the convolutional layer.

3.6 | LSTM and bi-directional LSTM models

LSTM is a particular artificial iterative neural network (RNN) architecture used in deep learning. It was suggested by Hochreiter et al.[69] It is proposed to overcome the long-term dependencies hijacking problem of the RNN architecture. Each input is calculated by taking into account the previous output value. The repeating modules in standard RNNs have a simple structure. However, this is more complicated for LSTM. While RNNs fail when the gap length increases, LSTM performs effectively as it can store data for a long time.[70] Unlike standard RNNs, it has feedback links. LSTMs, which can store data for a long time, are suitable for classifying, processing, and estimating time series.

In the LSTM architecture, which consists of repetitive sequential blocks that can determine the flow, there are three or four gates: "Forget", "Input", and "Output". These gates calculate a value between 0 and 1 using the logistic function to control the flow of information. The "Input Layer" controls whether the new input can be kept in memory by adding information useful to the cell during the learning phase. "Forget Layer" determines how long certain values can be kept in memory and clears out useless information during the learning phase. Finally, the "Output Layer" controls how much of the value in memory is used to calculate the output activation of the block to extract useful information from the cells. Figure 6 shows the LSTM architecture.

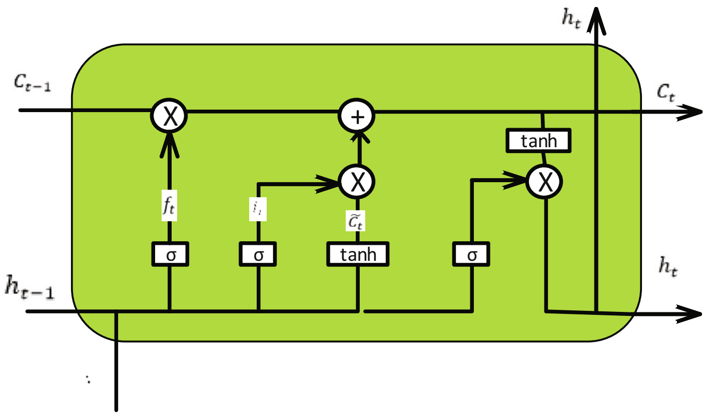


FIGURE 6 The basic architecture of long short-term memory (LSTM) neural network

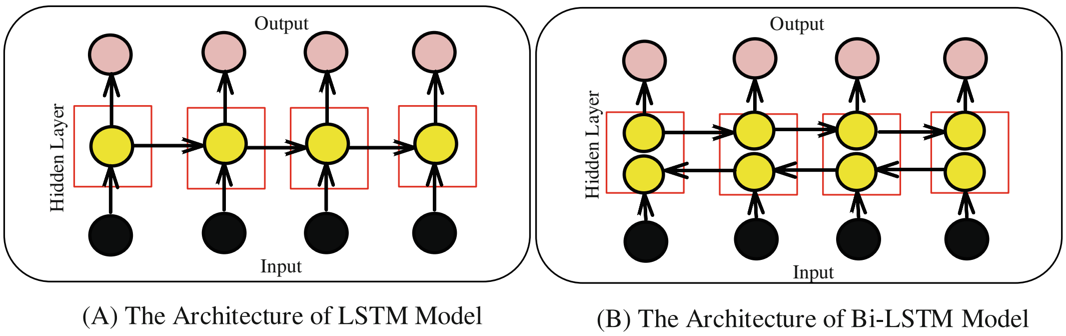


FIGURE 7 The architecture of (A) long short-term memory (LSTM) and (B) bidirectional long short-term memory (Bi-LSTM) model

BiLSTM is a bidirectional LSTM model proposed by Graves et al.[71] Figure 7 shows the general structure of the LSTM and Bi-LSTM models. It is a very useful approach in sort modeling as each hidden layer in the LSTM model has an input based on the calculation of the cell that processed the data at the previous time.[70] However, since the information can only be propagated forwardly in the LSTM model, the state of time *t* depends only on the previous information before time *t*. This may result in information loss. Unlike LSTM, the Bi-LSTM model has bidirectional information flow. This model handles all inputs equally. In this model, the array is trained by two LSTM networks. One of the LSTM networks trains in the forward direction while the other in the reverse direction. In the result layer, each LSTM network is connected to the same output layer.

3.7 | The proposed TCNN-BiLSTM model

In this study, a novel TCNN-BiLSTM deep learning model was proposed to detect rumors about the COVID-19 vaccines posted on Twitter. The proposed new model connects a two-stage convolutional neural network (TCNN) to the Bi-LSTM. Figure 8 shows the structure of the TCNN–Bi-LSTM model. In the proposed model, firstly, the collected raw dataset goes through the preprocessing steps. Then, significant information is extracted from textual tweets thanks to the FastText word embedding approach. Next, high-level features are captured from the input representation thanks to the two-stage TCNN model. Lastly, contextual information is extracted from the two-stage CNN layers using the Bi-LSTM layer to perform the sentiment analysis process.

4 | RESULTS

In this study, many experiments were conducted to prove the success of the proposed model. First, the performance of the proposed model was compared with baseline deep learning classification models, including CNN, LSTM, Bi-LSTM, CNN–LSTM, CNN–Bi-LSTM,

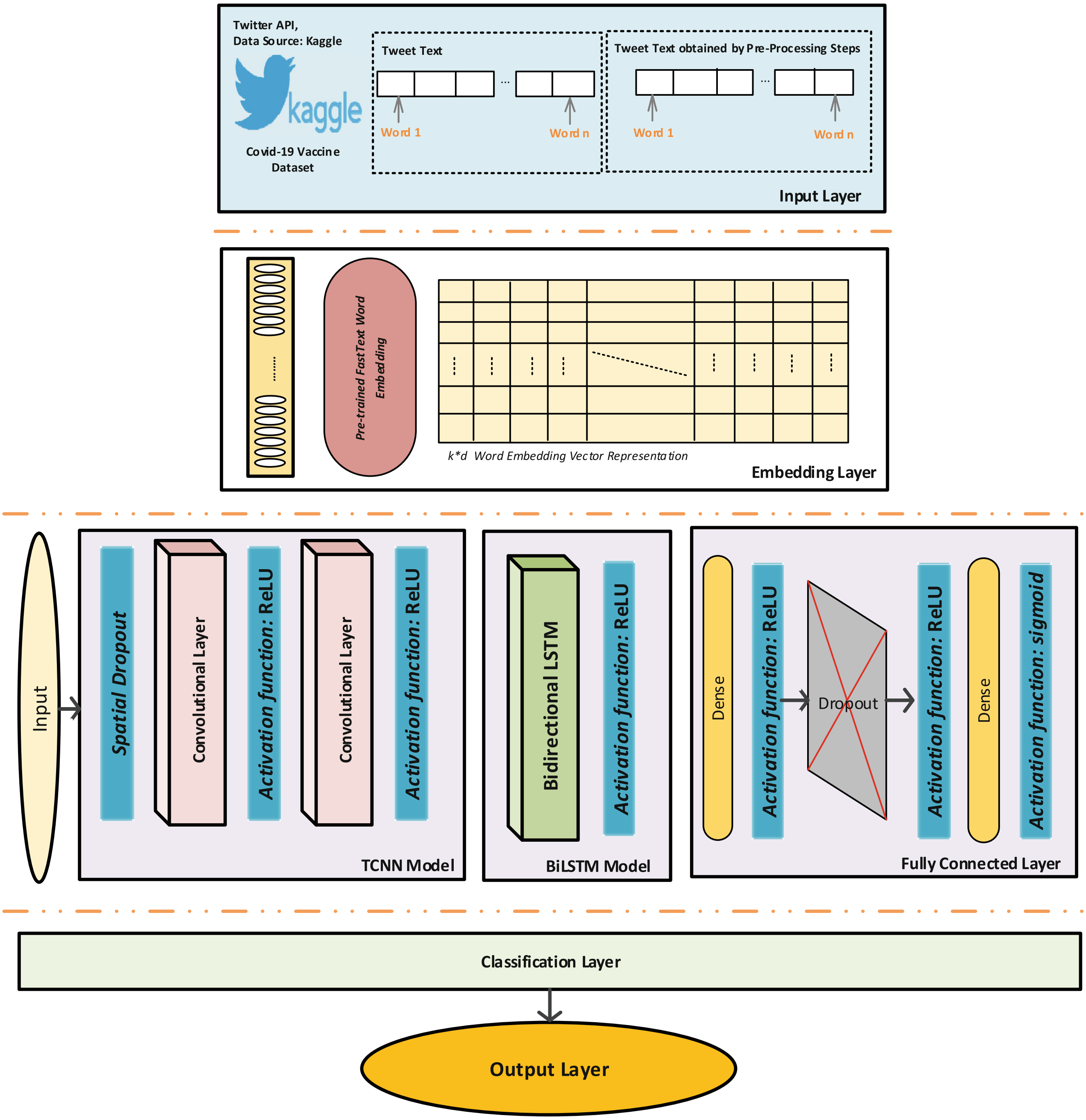


FIGURE 8 Structure of the proposed two-stage convolutional neural network (TCNN)–bidirectional long short-term memory (Bi-LSTM) model

LSTM-CNN, and Bi-LSTM–CNN. Then, the proposed model and compared deep learning models were trained with Glove and TFIDF word embedding techniques. And then, they were compared with FastText experimental results to demonstrate FastText's superior performance.

The performance of the proposed TCNN–Bi-LSTM model was also compared with seven machine learning classification models: Multinomial Naive Bayes,[33] Support Machine Classifier,[72] Random Forest,[73] Decision Tree,[74] K-Nearest Neighbors,[58] and Logistic Regression.[75] Experiments were tested on the Google Collaborate Pro platform using the Python programming language. Pandas, Keras, Numpy, spaCy, and Sklearn python programming libraries were used in the experiments.

4.1 | Evaluation metrics

The performances of the proposed classification model and the compared models were tested with four different commonly used evaluation metrics, which are "*Accuracy, Precision, Recall, and F1-score*". Also, both the training and test validation accuracy-loss curves were used to verify the performance of the proposed model.

Accuracy refers to the accuracy of the prediction and is calculated as follows:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{10}$$

The accuracy of the binary classification is calculated as positive and negative as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \tag{11}$$

In Equation (11), TP, TN, FP, and FN denote the true positive, true negative, false positive, and false negative.[57]

Precision describes the precision of a classifier and indicates the percentage of all clusters that are labeled as positive and that are indeed positive. It is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

Recall usually refers to a measure of integrity and indicates the percentage of true positive predictions that are correctly labeled. It is calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

The accuracy evaluation measure may not be a good metric due to the lack of a balanced dataset. In such cases, the F1-score is used. It is because the F1-score provides the results according to each target class. It is a statistical classification analysis criterion that takes into account both the precision of the classifier and the recall criteria.[76] It is calculated as follows:

$$\text{F1} - \text{score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

4.2 | Hyper parameter setting of the proposed and compared models

Table 2 presents the details of the parameters of the proposed model. This study adopts the hyper parameters to compare the proposed model with baseline deep learning models while performing the experimental results.

4.3 | Experimental results and analysis

In experiments, TF-IDF and Glove word embedding techniques were applied to demonstrate the performance of the FastText word embedding technique that was selected baseline in the proposed TCNN–Bi-LSTM model. K-fold cross validation technique with $k = 10$ was used to enhance the accuracy of the proposed model and other compared classifiers with proposed model in the experiments. Cross validation first divides the dataset

TABLE 2 Hyper parameter setting

Hyper parameter	
Platform	Google Colab
Optimizer	Adam
Loss function	Binary cross entropy
Learning rate	0.0001
Resampling technique	10-fold cross validation
SpatialDropout1D	0,3
Maximum length	70
Embedding dimension	FastText = 300 GloVe = 100
Batch size	32
Epoch	60
CNN filter size	100
CNN kernel size	3
CNN activation	ReLu
CNN maxpooling size	2
LSTM node	100
Bi-LSTM node	100
Dense 1	50
Dropout	0.25
Dense 2	1
Dense activation	Sigmoid

Abbreviations: Bi-LSTM, bidirectional long short-term memory; CNN, convolutional neural network; LSTM, long short-term memory.

into k equal subsets. Then, 1 subset of them is reserved as a test and k−1 subsets are used for training. This process is repeated k times in total. The result validation value is the average value across all k experiments.

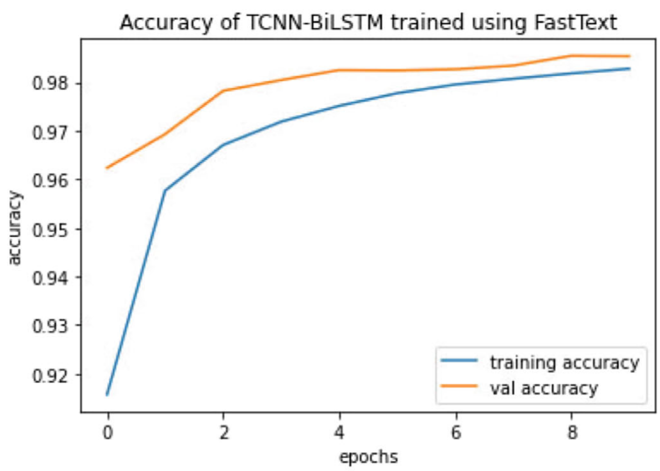
Table 3 shows the performance comparison of the proposed model with other deep learning models using three word-embedding vectors. The experimental results compared the proposed model with all variations of CNN, LSTM, and Bi-LSTM deep learning models. Figures 9, 10, and 11 show the proposed model's accuracy and loss curve for training and test data with the FastText, Glove, and TF-IDF word embedding vectors. Additionally, the performance of the proposed TCNN–Bi-LSTM model was compared with seven machine learning classification models with FastText and TF-IDF word embedding vectors. Table 4 shows the performance comparison of the proposed model with basic machine learning classifiers.

As seen in Table 3, the proposed TCNN–Bi-LSTM model with the FastText word embedding technique has a higher performance than all other baseline deep learning models in the experiments on the collected dataset. However, the TCNN–Bi-LSTM model outperforms the CNN–Bi-LSTM model. These results reveal that the two-stage convolutional layer makes learning more effective. As indicated by bold values in Table 3, the proposed TCNN-BiLSTM model has gained the best performance in all compared word embedding techniques (FastText, Glove, and TF-IDF) for all evaluation criteria. A very successful learning model has been created since the output of the two-stage TCNN model used to extract stronger local features feeds the Bi-LSTM model that recalls forward-looking information. FastText is pre-trained with location-based and sub-word information features; it can process hidden features in language and out-of-vocabulary words better than other models. Experimental results have proven that word vector generation with the FastText word embedding technique gives better results than other approaches such as Glove and TF-IDF. This proves that the word embedding vector technique used in the proposed model extracts stronger features in sentiment analysis. While the LSTM architecture only performs forward transitions, the Bi-LSTM model performs forward and backward transitions (See Figure 7). Moreover, as seen in Table 3, Bi-LSTM models performed better than LSTM models when we applied both LSTM and Bi-LSTM classifier models to our dataset. In this way, the network performs more successful deep learning using Bi-LSTM. Moreover, the experimental results of the combination models of CNN–LSTM and CNN–Bi-LSTM prove the power of Bi-LSTM.

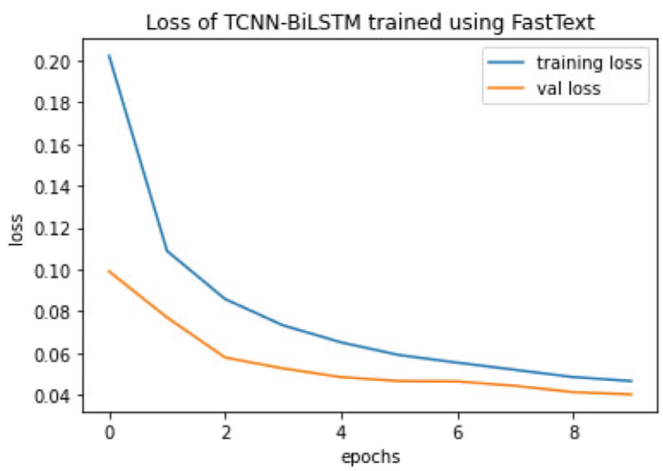
TABLE 3 Performance comparison with baseline deep learning models

Embedding model	Deep learning classification model	Accuracy	Precision	Recall	F1-Score
FastText	TCNN-BİLSTM	**0.963090**	**0.977657**	**0.972550**	**0.975053**
	CNN	0.948183	0.956750	0.962725	0.964691
	LSTM	0.946098	0.951033	0.947701	0.964074
	Bi-LSTM	0.947130	0.958540	0.957881	0.964473
	CNN–LSTM	0.959194	0.969788	0.962644	0.967382
	CNN–Bi-LSTM	0.959484	0.961270	0.964440	0.972782
	LSTM–CNN	0.950043	0.9638928	0.9683766	0.966068
	Bi-LSTM–CNN	0.951809	0.9663544	0.9685028	0.967339
Glove	**TCNN–Bi-LSTM**	**0.942166**	**0.962646**	**0.954641**	**0.9610865**
	CNN	0.892178	0.914038	0.936709	0.955235
	LSTM	0.917714	0.962091	0.928917	0.945213
	Bi-LSTM	0.931377	0.954019	0.9360918	0.945937
	CNN–LSTM	0.931030	0.959767	0.953805	0.956777
	CNN–Bi-LSTM	0.933209	0.959252	0.950725	0.957970
	LSTM–CNN	0.891324	0.918420	0.933502	0.925794
	Bi-LSTM–CNN	0.893138	0.941474	0.916091	0.928585
TF-IDF	**TCNN–Bi-LSTM**	**0.885903**	**0.9495505**	**0.901322**	**0.924740**
	CNN	0.862473	0.899311	0.883922	0.885037
	LSTM	0.720221	0.758760	0.726667	0.737984
	Bi-LSTM	0.737545	0.785378	0.737405	0.748761
	CNN–LSTM	0.851454	0.864080	0.861155	0.870896
	CNN–Bi-LSTM	0.858375	0.872813	0.878453	0.864984
	LSTM–CNN	0.851960	0.861205	0.865918	0.876496
	Bi-LSTM–CNN	0.868828	0.891297	0.883975	0.887870

Abbreviations: Bi-LSTM, bidirectional long short-term memory; CNN, convolutional neural network; LSTM, long short-term memory; TCNN, two-stage convolutional neural network.

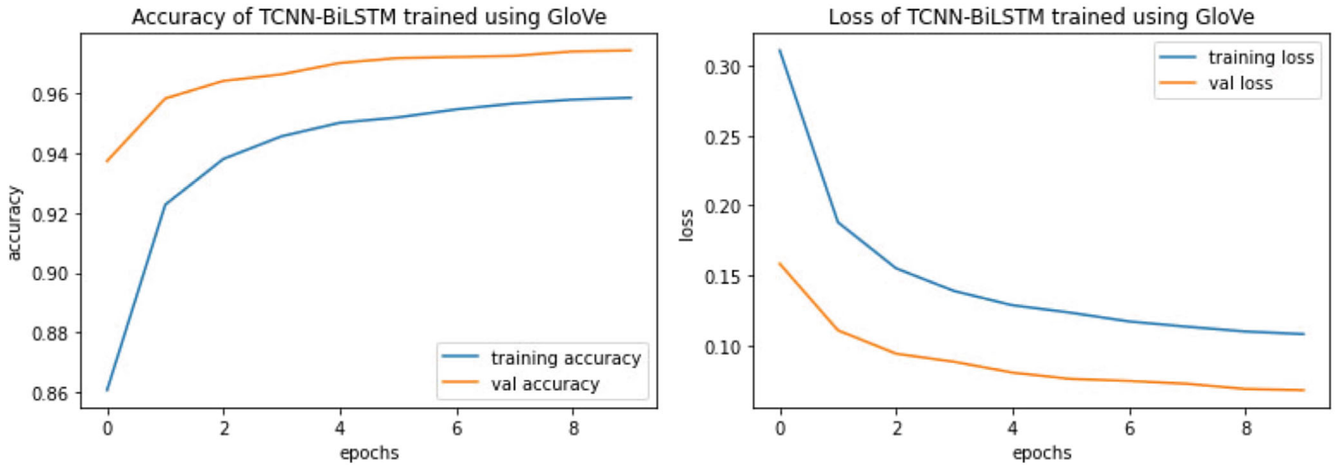


(A)



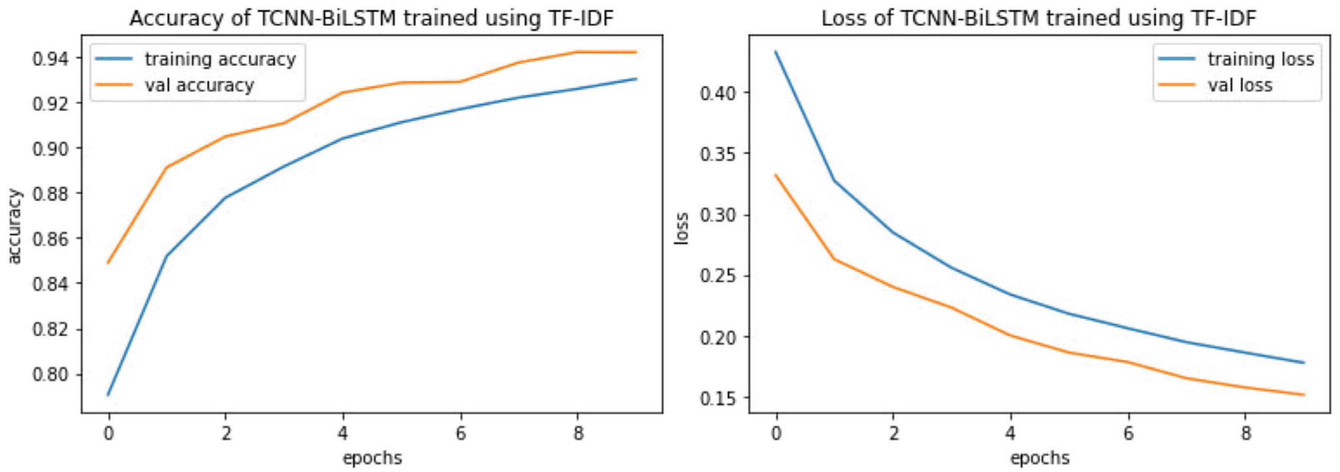
(B)

FIGURE 9 (A) Accuracy of two-stage convolutional neural network (TCNN)–bidirectional long short-term memory (Bi-LSTM) model trained using FastText, (B) Loss of TCNN–Bi-LSTM model trained using FastText



(A) (B)

FIGURE 10 (A) Accuracy of two-stage convolutional neural network (TCNN)–bidirectional long short-term memory (Bi-LSTM) model trained using GloVe, (B) Loss of TCNN–Bi-LSTM model trained using GloVe



(A) (B)

FIGURE 11 (A) Accuracy of two-stage convolutional neural network (TCNN)–bidirectional long short-term memory (Bi-LSTM) model trained using TF-IDF, (B) Loss of TCNN–Bi-LSTM model trained using TF-IDF

The accuracy and loss curves of the proposed TCNN–Bi-LSTM model with FastText, Glove, and TF-IDF are shown in Figures 9, 10, and 11, respectively. As shown in the figures, the performance scores of the proposed model in all word vectors are higher than the other models. The accuracy was calculated at 0.98% for the proposed model with FastText. As shown in Figures 9, 10, and 11 the training accuracy of the proposed model with FastText, Glove, and TF-IDF started at approximately 89%, 82%, and 78%, respectively. Then they ended with 97%, 95%, and 93%, respectively.

Table 4, the performance comparison of the proposed model with basic machine learning classifiers. As indicated by bold values in Table 4, the proposed TCNN-BiLSTM model has outperformed the compared machine learning models in both FastText and TF-IDF word embedding techniques for all evaluation criteria. Also, the average accuracy scores of deep learning architectures are considerably higher than machine learning architectures. The most significant reason for this is that the learning process is divided into small steps, and the results of each step are combined into a single output in machine learning methods. In contrast, the learning process is processed from end to end in deep learning methods. Moreover, machine-learning methods need to be taught when the result is wrong. In contrast, deep learning methods place the data in a hierarchy of different concepts and provide learning from their own mistakes thanks to the multi-level layers of neural networks. Additionally, Table 4 shows that while the FastText data embedding technique achieves the best performance compared to other word embedding techniques in all deep learning classification models, it has not had the same outstanding success in machine learning models.

TABLE 4 Performance comparison of the proposed model with machine learning classifiers

Embedding model	Classification model	Accuracy	Precision	Recall	F1-Score
FastText	**TCNN-BİLSTM**	**0.963090**	**0.977657**	**0.972550**	**0.975053**
	SVM	0.7178	0.7092	0.7265	0.7012
	Multinomial NB	0.7917	0.8036	0.7984	0.7905
	LR	0.7349	0.7321	0.7341	0.7507
	RF	0.8381	0.8207	0.8265	0.8349
	DT	0.7725	0.7812	0.7796	0.7791
	KNN	0.7680	0.7604	0.7725	0.7510
TF-IDF	**TCNN–BİLSTM**	**0.885903**	**0.9495505**	**0.901322**	**0.924740**
	SVM	0.7652	0.7692	0.7853	0.7644
	Multinomial NB	0.7879	0.7783	0.7912	0.7746
	LR	0.8557	0.8618	0.8788	0.8702
	RF	0.8359	0.8216	0.8263	0.8283
	DT	0.8445	0.8523	0.8451	0.8486
	KNN	0.8172	0.8341	0.9389	0.8834

Abbreviations: Bi-LSTM, bidirectional long short-term memory; TCNN, two-stage convolutional neural network.

5 | CONCLUSIONS

This research aims to investigate the sentiments of tweets about the COVID-19 vaccines and comments on these tweets among Twitter users by using the power of Twitter data. The proposed model is a Two-stage CNN (TCNN) and Bidirectional LSTM (Bi-LSTM) architecture with FastText word embedding technique. Due to FastText being pre-trained with location-based and sub-word information features, it can process hidden features in language and out-of-vocabulary words. Thanks to FastText, strong features could be extracted in sentiment analysis. Experimental results prove that the proposed model achieves the highest performance by using FastText. In the proposed TCNN model, the output of the first convolutional layer feeds the next convolutional layer to maximize the advantages of the CNN model. In this way, stronger local features were extracted from the pre-processed tweets. Then, the output of the TCNN model was sent as input to the Bi-LSTM model to capture long-distance dependencies, and these features were combined into a single hybrid TCNN–Bi-LSTM model proposed. The two-stage TCNN model captures and learns local information sufficiently in the proposed novel model, while the Bi-LSTM model remembers forward-looking information.

The experiments were conducted on a new dataset collected using MaxQDA qualitative data analysis tool via the Twitter API platform. The proposed model was tested with machine learning and deep learning classifier models. Also, all models were trained using the TF-IDF and Glove word embedding techniques to prove the performance of the word embedding technique used in the model proposed in the study. As a result, the proposed model outperformed compared to both deep learning and machine learning classifiers.

The experimental results of this study will help healthcare researchers to collect accurate information about the vaccination process. Also, it will facilitate detecting inappropriate, incomplete, and erroneous information about vaccination. The results of this study will enable society to broaden its perspective on the administered vaccines. It can also assist the government and healthcare agencies to plan and implement the promotion of the vaccination in a timely manner to achieve the herd immunity provided by the vaccination.

CONFLICT OF INTEREST

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

If someone wants to request the data, they should contact Serpil Aslan at “serpil.aslan@ozal.edu.tr” mail address. Data available on request from the authors.

ORCID

Serpil Aslan https://orcid.org/0000-0001-8009-063X

REFERENCES

1. World Health Organization. Mental Health and Psychosocial Considerations during the COVID-19 Outbreak. Available online: 11 March 2022.
2. Worldometer. Available online: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1? (Accessed: 11 March 2022).
3. Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien AE. Sentiment analysis of COVID-19 tweets by deep learning classifiers—a study to show how popularity is affecting accuracy in social media. *Appl Soft Comput*. 2020;97:106754.
4. Rehman M, Tauseef I, Aalia B, Shah SH, Junaid M, Haleem KS. Therapeutic and vaccine strategies against SARS-CoV-2: past, present and future. *Future Virol*. 2020;15(7):471-482.
5. Cabanillas B, Novak N. Allergy to COVID-19 vaccines: a current update. *Allergol Int*. 2021;70(3):313-318.
6. Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10).
7. Mahdikhani M. Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of Covid-19 pandemic. *Int J Inf Manag Data Insights*. 2022;2(1):100053.
8. Okango E, Mwambi H. Dictionary based global twitter sentiment analysis of coronavirus (COVID-19) effects and response. *Ann Data Sci*. 2022;9(1):175-186.
9. Valle-Cruz D, Fernandez-Cortez V, López-Chau A, Sandoval-Almazán R. Does twitter affect stock market decisions? Financial sentiment analysis during pandemics: a comparative study of the H1N1 and the COVID-19 periods. *Cogn Comput*. 2022;14(1):372-387.
10. Chinnasamy P, Suresh V, Ramprathap K, Jebamani BJA, Srinivas Rao K, Shiva Kranthi M. COVID-19 vaccine sentiment analysis using public opinions on twitter. *Mater Today: Proc*. 2022;64:448–451.
11. Nielsen, F. Å. (2011). A new ANEW: evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.
12. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. J. (2011). Sentiment analysis of twitter data. In Proceedings of the Workshop on Language in Social Media (LSM 2011) (pp. 30–38).
13. Pang, B., & Lee, L. (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *arXiv preprint cs/0409058*.
14. Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*.
15. Chiha R, Ayed MB, Pereira CDC. A complete framework for aspect-level and sentence-level sentiment analysis. *Appl Intell*. 2022;1-19.
16. Kim, S. M., & Hovy, E. (2004). Determining the sentiment of opinions. In COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics (pp. 1367–1373).
17. Aygün İ, Kaya B, Kaya M. Aspect based twitter sentiment analysis on vaccination and vaccine types in covid-19 pandemic with deep learning. *IEEE J Biomed Health Inform*. 2021;26(5):2360-2369.
18. Agarwal, A., Biadsy, F., & Mckeown, K. (2009). Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009) (pp. 24–32).
19. Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The Development and Psychometric Properties of LIWC2015.
20. Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 168–177).
21. Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): instruction manual and affective ratings (Vol. 30, No. 1, pp. 25–36). *Technical Report C-1, the Center for Research in Psychophysiology*, University of Florida.
22. Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10).
23. Cambria, E., Havasi, C., & Hussain, A. (2012). Senticnet 2: a semantic and affective resource for opinion mining and sentiment analysis. In Twenty-Fifth International FLAIRS Conference.
24. Chen, N., & Wang, P. (2018). Advanced combined lstm-cnn model for twitter sentiment analysis. In 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS) (pp. 684–687). IEEE.
25. Xia R, Zong C, Li S. Ensemble of feature sets and classification algorithms for sentiment classification. *Inform Sci*. 2011;181(6):1138-1152.
26. Zhou S, Chen Q, Wang X. Fuzzy deep belief networks for semi-supervised sentiment classification. *Neurocomput*. 2014;131:312-322.
27. Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A convolutional neural network for modelling sentences. arXiv preprint arXiv: 1404.2188.
28. Irsoy, O., & Cardie, C. (2014). Deep recursive neural networks for compositionality in language. In Advances in Neural Information Processing Systems (pp. 2096-2104).
29. Tai, K. S., Socher, R., & Manning, C. D. (2015). Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*.
30. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Trans Signal Process*. 1997;45(11):2673-2681.
31. Kumar A, Sebastian TM. Sentiment analysis on twitter. *Int J Comput Sci Issues (IJCSI)*. 2012;9(4):372.
32. Joshi, R., & Tekchandani, R. (2016). Comparative analysis of twitter data using supervised classifiers. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 3, pp. 1–6). IEEE.
33. Parikh R, Movassate M. Sentiment analysis of user-generated twitter updates using various classification techniques. *CS224N Final Report*. 2009;118:1-18.
34. Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*. 2009;1(12):2009.
35. Arun C, Karthick S, Selvakumarasamy S, James SJ. Car parking location tracking, routing and occupancy monitoring system using cloud infrastructure. *Mater Today: Proc*. 2021.
36. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems (pp. 3111–3119).

37. Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1532–1543).
38. Rubtsova Y. Automatic term extraction for sentiment classification of dynamically updated text collections into three classes. *International Conference on Knowledge Engineering and the Semantic Web*. Springer; 2014:140-149.
39. Chen, X., Tang, W., Xu, H., & Hu, X. (2014). Double Lda: a sentiment analysis model based on topic model. In 2014 10th International Conference on Semantics, Knowledge and Grids (pp. 49–56). IEEE.
40. Fu X, Liu W, Xu Y, Cui L. Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. *Neurocomput*. 2017;241:18-27.
41. Abid F, Alam M, Yasir M, Li C. Sentiment analysis through recurrent variants latterly on convolutional neural network of twitter. *Future Gener Comput Syst*. 2019;95:292-308.
42. Araque O, Corcuera-Platas I, Sánchez-Rada JF, Iglesias CA. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Syst Appl*. 2017;77:236-246.
43. Zhao R, Mao K. Fuzzy bag-of-words model for document representation. *IEEE Trans Fuzzy Syst*. 2017;26(2):794-804.
44. Yoon, J., & Kim, H. (2017). Multi-channel lexicon integrated CNN-BiLSTM models for sentiment analysis. In Proceedings of the 29th Conference on Computational Linguistics and Speech Processing (ROCLING 2017) (pp. 244–253).
45. Al-Sarem M, Alsaeedi A, Saeed F, Boulila W, AmeerBakhsh O. A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs. *Appl Sci*. 2021;11(17):7940.
46. Kamyab M, Liu G, Adjeisah M. Attention-based CNN and Bi-LSTM model based on TF-IDF and GloVe word embedding for sentiment analysis. *Appl Sci*. 2021;11(23):11255.
47. Barkur G, Vibha GBK. Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: evidence from India. *Asian J Psychiatr*. 2020;51:102089.
48. Nwani CD, Lakra WS, Nagpure NS, Kumar R, Kushwaha B, Srivastava SK. Toxicity of the herbicide atrazine: effects on lipid peroxidation and activities of antioxidant enzymes in the freshwater fish Channa punctatus (Bloch). *Int J Environ Res Public Health*. 2010;7(8):3298-3312.
49. Süral I, Griffiths MD, Kircaburun K, Emirtekin E. Trait emotional intelligence and problematic social media use among adults: the mediating role of social media use motives. *Int J Ment Health Addict*. 2019;17(2):336-345.
50. Xue J., Chen J., Hu R., Chen C., Zheng C., Liu X., Zhu T., Twitter discussions and emotions about COVID-19 pandemic: a machine learning approach (2020). arXiv: 2005.12830.
51. Feng, Y., & Zhou, W. (2020). Is working from home the new norm? An observational study based on a large geo-tagged COVID-19 twitter dataset. *arXiv preprint arXiv:2006.08581*.
52. Hornung, O. et al. (2018). When emotions go social–understanding the role of emotional intelligence in social network use. In Proceedings of Twenty-Sixth European Conference on Information Systems, pp. 40.
53. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. *J Travel Med*. 2020;27(3):taaa031.
54. Jelodar H, Wang Y, Orji R, Huang S. Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach. *IEEE J Biomed Health Inform*. 2020;24(10):2733-2742.
55. Pano T, Kashef R. A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. *Big Data Cogn Comput*. 2020;4(4):33.
56. Ortiz-Sánchez E, Velando-Soriano A, Pradas-Hernández L, Vargas-Román K, Gómez-Urquiza JL, Albendín-García L. Analysis of the anti-vaccine movement in social networks: a systematic review. *Int J Environ Res Public Health*. 2020;17(15):5394.
57. Praveen SV, Ittamalla R, Deepak G. Analyzing the attitude of Indian citizens towards COVID-19 vaccine–a text analytics study. *Diabetes Metab Syndr Clin Res Rev*. 2021;15(2):595-599.
58. Shamrat FMJM, Chakraborty S, Imran MM, et al. Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. *Indones J Electr Eng Comput Sci*. 2021;23(1):463-470.
59. Alam KN, Khan MS, Dhruba AR, et al. Deep learning-based sentiment analysis of COVID-19 vaccination responses from Twitter data. *Comput Math Methods Med*. 2021;2021(2021):1-15.
60. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on twitter. *Int J Infect Dis*. 2021;108:256-262.
61. Investor fact sheet. Twitter. 2021. [29-04-2021] https://s22.q4cdn.com/826641620/files/doc_financials/2021/q1/Q1'21-Shareholder-Letter.pdf.
62. We used MAXQDA 2020 (VERBI Software, 2019) for data analysis.
63. Loria, S. (2018). textblob Documentation. Release 0.15, 2, 269.
64. Saleena N. An ensemble classification system for twitter sentiment analysis. *Procedia Comput Sci*. 2018;132:937-946.
65. Afroz, N., Boral, M., Sharma, V., & Gupta, M. (2021). Sentiment analysis of COVID-19 nationwide lockdown effect in India. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 561–567). IEEE.
66. Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., & Joulin, A. (2017). Advances in pre-training distributed word representations. *arXiv preprint arXiv:1712.09405*.
67. Priyadarshini I, Puri V. A convolutional neural network (CNN) based ensemble model for exoplanet detection. *Earth Sci Inform*. 2021;14(2):735-747.
68. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.
69. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735-1780.
70. Mohan, A. T., & Gaitonde, D. V. (2018). A deep learning based approach to reduced order modeling for turbulent flow control using LSTM neural networks. *arXiv preprint arXiv:1804.09269*.
71. Graves A, Fernández S, Schmidhuber J. Bidirectional LSTM networks for improved phoneme classification and recognition. *International Conference on Artificial Neural Networks*. Springer; 2005:799-804.
72. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297.
73. Da Silva NF, Hruschka ER, Hruschka ER Jr. Tweet sentiment analysis with classifier ensembles. *Decis Support Syst*. 2014;66:170-179.
74. Fitri VA, Andreswari R, Hasibuan MA. Sentiment analysis of social media twitter with case of anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Comput Sci*. 2019;161:765-772.

75. Lin, J., & Kolcz, A. (2012). Large-scale machine learning at twitter. In Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data (pp. 793–804).
76. Aslan S, Kaya M. Topic recommendation for authors as a link prediction problem. *Future Gener Comput Syst*. 2018;89:249-264.

How to cite this article: Aslan S. A novel TCNN–Bi-LSTM deep learning model for predicting sentiments of tweets about COVID-19 vaccines. *Concurrency Computat Pract Exper*. 2022;34(28):e7387. doi: 10.1002/cpe.7387