New Generation Computing (2024) 42:89–107
https://doi.org/10.1007/s00354-023-00227-0

A CNN-LSTM-Based Hybrid Deep Learning Approach for Sentiment Analysis on Monkeypox Tweets

Krishna Kumar Mohbey[1] · Gaurav Meena[1] · Sunil Kumar[2] · K. Lokesh[1]

Received: 4 September 2022 / Accepted: 21 July 2023 / Published online: 14 August 2023

Abstract
The research on sentiment analysis has shown a great deal of utility in the field of public health, specifically in the investigation of infectious illnesses. As the world begins to recuperate from the devastating effects of the COVID-19 pandemic, there is a growing concern that a different pandemic, known as Monkeypox, may strike the world once more. The contagious illness known as Monkeypox has been documented in over 73 countries worldwide. This unexpected epidemic has become a significant cause of anxiety for many people and health authorities. Various social media platforms have presented various perspectives regarding the monkeypox epidemic. Our goal is to research how the public feels about the recent Monkeypox epidemic to assist policymakers in developing a deeper comprehension of how the public views the illness. This research uses a CNN-LSTM-based hybrid architecture to ascertain people's feelings regarding Monkeypox disease. A series of experiments were conducted on an open-access dataset of tweets related to the Monkeypox. The tweets undergo various pre-processing, global vectorization, and one-hot encoding techniques. According to the findings of our experiments, the hybrid model provided better accuracy, which was approximately 91%. In addition, the findings are validated by contrasting them with more conventional machine learning techniques. The outcomes of this investigation contribute to a general population that has a greater awareness of the Monkeypox infection.

Keywords Monkeypox · COVID-19 · LSTM · Deep learning · Sentiment polarities · Machine learning

✉ Gaurav Meena
gaurav.meena@curaj.ac.in

[1] Department of Computer Science, Central University of Rajasthan, Ajmer, India

[2] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

1 Introduction

The clinical presentations of Monkeypox are comparatively milder than those of smallpox. The manifestation of this condition is characterized by symptoms that bear a resemblance to those exhibited by smallpox. Following the discontinuation of smallpox vaccinations and the declaration of smallpox eradication in 1980, Monkeypox has emerged as the foremost orthopoxvirus concerning public health concerns. Monkeypox is an infectious ailment that predominantly affects individuals in Western and Central Africa. Nevertheless, the species has recently extended its range to metropolitan regions and is frequently observed near equatorial rainforests. A diverse array of rodent and non-human primate species serves as hosts for various organisms. The monkeypox virus is classified as a member of the Poxviridae viral family, characterized by double-stranded DNA. Specifically, it is categorized within the Orthopoxvirus genus. The monkeypox species of viruses have been identified to have originated from the Congo Basin and western Africa. This observation indicates that these geographical locations may harbour genetically diverse subcategories of the virus. According to established beliefs, the Congo Basin clade was deemed to be the most contagious and responsible for inducing the most acute symptoms. The geopolitical demarcations of Cameroon function as a de facto partition between the two viral clades, given that it is the singular nation where both clades have been detected. The zoonotic transmission of the monkeypox virus to humans is primarily attributed to contact with infected animals, particularly rats. Although an epidemic has the potential to transmit from one individual to another, it is not a sustainable mode of propagation over an indefinite period. The clinical manifestations bear a resemblance to those of smallpox, albeit with less severity.

Researchers in science and psychology are dedicated to gaining a deeper understanding of the human experience, including the domains of psychology and mental health [33]. These scholars are particularly interested in the rapid proliferation of information facilitated by social media platforms. Twitter, among other social media platforms, has been used to collect data in behavioral science and psychology. This tool has also been used to determine a user's personality traits [1, 2] and gain insight into the browsing habits and past behaviors of internet users. Sentiment analysis can be utilized to categorize an individual's viewpoint into one of three classifications, namely positive, negative, or neutral, concerning a particular subject matter, commodity, film, or current event. The present system utilizes NLP techniques to automatically analyze and quantify users' emotions by examining Twitter data. Sentiment Analysis research concerns positive, negative, and neutral sentiments conveyed through tweets. The sentiment analysis task involves facile access to public opinions expressed in evaluations and survey responses. The capture of textual information, the acquisition of visual content, the production of motion pictures, and the recognition of auditory signals are among the potential uses of this technology. After data collection, the input is dissected into its constituent parts, which may consist of discrete words or phrases, to conduct sentiment analysis. This analytical process is conducted on data that is analogous to those tweets. Using sentiment analysis on social media platforms like Twitter [3] has increased the demand for public perspectives. These viewpoints are gathered in textual format. Accurately forecasting data from tweets using text analysis to meet commercial evaluation requirements is a complex challenge. Twitter enables individuals to express their sentiments, ideas, and viewpoints regarding the worldwide pandemic [4].

In the last month of 2019, a new viral attack happened worldwide, which messed up everyone's lives. All the connected study groups are trying to figure out how a pandemic acts so they can figure out when it will end, but every time it gives them new values for different parameters, it surprises them [34, 35]. Despite the ongoing pandemic, individuals persist in utilizing Twitter [31]. Utilizing machine learning algorithms is necessary considering the difficulties associated with determining the intrinsic significance of content through natural language processing (NLP) techniques, which include contextual phrases and words, as well as the ambiguity present in written or spoken language [5–7]. The field of NLP has significantly influenced the analysis of textual data and the classification of social media content. Transfer learning refers to acquiring knowledge or skills that can be applied in diverse settings or situations beyond the original context in which they were learned. The utilization of transfer learning is a common practice in image processing, aiming to expedite the learning and training procedures. Knowledge transfer plays a crucial role in the transference of acquired knowledge [8] from the original field to the intended field. As per the research conducted by Sv et al. [22], it has been observed that the sentiment on social media regarding the monkeypox virus is predominantly positive (28.82%) as compared to negative (23.01%). Upon closer examination, it becomes apparent that most tweets expressing positive sentiments regarding Monkeypox pertain to the relatively mild symptoms and lower mortality rate associated with the infection. According to a public opinion survey, individuals do not appear overly concerned about the monkeypox virus. The examination of tweets expressing negative sentiments towards Monkeypox unveiled that users were deliberating on various topics such as the virus's fatality potential, its severity, the resulting lesions, vaccine accessibility, the possibility of Monkeypox becoming the next pandemic after COVID-19, the safety of travel, and the implications of the virus's transmission on human health. Presented here is an itemization of the technical components that we provided.

- The present study introduces a hybrid approach based on CNN and LSTM to identify users' sentiments on a monkeypox dataset.
- The tweets are processed using various techniques, including pre-processing, glob vectorization, and one hot encoding.
- The efficacy of the proposed methodology is confirmed through comparison with several conventional machine learning techniques.
- Monkeypox infections could be brought to public attention by disseminating information based on identified polarities within society.

The rest of the paper is presented in the following manner. In part 2, the related studies are summarized.

Part 3 discusses the proposed hybrid architecture in detail. Evaluation findings and assessments are included in part 4. Finally, part 5 concludes with a consideration of future work.

2 Related Works

Monkeypox is a matter of significant concern for countries in West and Central Africa and the global community due to its widespread impact on public health. In 2003, the United States of America encountered an outbreak of Monkeypox, marking the first instance of the disease occurring outside of Africa. Epidemiologists identified the disease's etiology as being linked to the exposure of captive individuals to infected prairie dogs. The animals above cohabited in a confined space alongside dormice and Gambian pouched rats that were brought into Ghana from elsewhere. The current outbreak has led to the documentation of more than 70 cases of Monkeypox in the United States. According to reports, Nigerian tourists contracting Monkeypox were documented in various countries, including Israel in September 2018, the UK in September 2018, December 2019, May 2021, and May 2022, Singapore in May 2019, and the USA in July and November 2021. The countries mentioned above are in the region of Southeast Asia. In May 2022, several instances of Monkeypox were detected in nations where the disease was not typically prevalent. Current research endeavors are underway to expand the knowledge base surrounding the ailment above's epidemiology, vectors, and transmission dynamics [9]. The orthopoxvirus, known as Monkeypox, can potentially infect humans and cause a viral illness characterised by symptoms such as fever and rash, which resemble smallpox. Following the eradication, the smallpox virus from the human population in 1980, Monkeypox has emerged as the most severe orthopoxvirus infection in humans. The disease is frequently documented in rural regions of Central and West African countries, particularly near tropical rainforests where individuals may come into contact with infected animals. Monkeypox can be transmitted through direct contact with respiratory droplets of an infected individual, either in a domestic or medical setting or through exposure to contaminated materials such as bedding. Despite being the primary mode of person-to-person transmission, monkeypox outbreaks frequently manifest in little clusters of a few cases without advancing to extensive community transmission. This is due to the high level of contagion associated with Monkeypox. Prompt intervention measures facilitate the rapid containment of a disease outbreak. Monkeypox has been documented in multiple nations as a result of either the introduction of the disease by travelers or animals that have been infected [10]. On May 20, 2022, the World Health Organization (WHO) organized an urgent conference to address global concerns regarding the escalating incidence of the monkeypox virus. The meeting was convened to discuss worldwide apprehensions [11]. The WHO engaged in a period of contemplation spanning several days to determine whether the epidemic in question met the criteria for classification as a "potential public health emergency of international concern" (PHEIC), a designation previously applied to outbreaks of COVID-19 and Ebola [12]. On June 6, 2022, the Center for Disease Control (CDC) in the United States released a notification about

Monkeypox, categorized as "Level 2". This statement was made in reaction to the notable increase in documented instances [13]. The WHO declared Monkeypox a global health emergency after a conference on July 23, 2022.

A generalized prediction model based on an artificial neural network (ANN) model was created by Kuvvetli et al. [14] to match the distributions of various nations and forecast the future number of daily cases and fatalities by COVID-19. The researchers utilized data that was gathered across multiple countries during the period spanning from March 11, 2020, to January 23, 2021. The utilization of the ANN model is a potential approach the government can employ to mitigate issues in hospitals and other healthcare facilities [32]. Opinion mining, also called emotional extraction, is a methodology utilized to examine customers' sentiments [15]. This technique can be executed by applying either Text Mining (TM) or NLP. Using sentiment analysis offers several benefits, such as increasing upselling opportunities, monitoring agents, and accessing real-time data. Service providers may utilize sentiment analysis to discern customer satisfaction levels and identify areas for improvement [16]. Identifying positive and negative sentiments within intricate wordplays can present a challenge. However, implementing sentiment analysis can aid in this task [17]. By employing the sentiment analysis methodology, scholars can ascertain the comprehensive mood of a written composition. Sentiment analysis enhances the process of gathering information by increasing its comprehensiveness.

Hassan et al. [18] proposed a method for polarity identification that integrates CNN and Long Short-Term Memory (LSTM) models, utilizing pre-trained word vectors sourced from IMDB movie reviews. The developed classifier utilizing this approach comprised a convolutional processing layer with four layers, two layers of pooling processing and two layers of output processing. The CNN and LSTM models exhibited lower accuracy rates of 87.0% and 81.8%, respectively, compared to the combined CNN + LSTM models, which achieved an accuracy rate of 88.3% during the trials. Shen et al. [19] proposed a novel methodology to ascertain the sentiment polarity of film critiques by incorporating CNN with Bidirectional LSTM. The approach above effectively identified and distinguished between affirmative and negative evaluations. The integration of CNN and LSTM classifiers yields a model with an accuracy of 89.7%, surpassing the accuracy of the individual models, which are 83.9% and 78.5%, respectively. The CNN classifier formulated using this approach comprised two stages of convolutional processing, two stages of pooling processing, and two strata of output processing. The combined utilization of CNN and LSTM models increased accuracy from 81.8 to 88.3% in the trials. Using the CNN approach is a viable method for conducting sentiment analysis on a collection of tweets. Previous research has demonstrated a more significant influence [20]. Text and images may be used to identify and categorize sentiments in many ways. Several well-known image datasets, such as JAFFE, are also available to do image-based sentiment analysis [23, 24].

The limitation of short-term memory poses a significant challenge for most Recurrent Neural Networks (RNNs), thereby necessitating the inclusion of LSTM as a crucial component of machine learning [21]. Consequently, the retention and application of data across different processes pose a challenge. As exemplified, in cases with a dataset or numerical data to be processed, RNNs can overlook certain historical information [22]. The field of infodemiology has been revolutionized by using data derived from social media platforms, which has substantially facilitated the examination of human-related phenomena. Furthermore, these social media platforms disseminate diverse statistical information about illness trends, such as Monkeypox, including the number of comments, images, and videos exchanged. Consequently, this enables the anticipation of monkeypox morbidity rates in various regions and notifies the responsible health policy decision-makers about the necessity of devising educational and preventive measures in the regions with the highest susceptibility [36].

3 CNN-LSTM-Based Sentiment Analysis

This section describes the sentiment classification system. The suggested design is built on the combination of CNN and LSTM. An open-access dataset of Tweets about Monkeypox that were posted on Twitter from 7 May 2022 to 9 October 2022 [25] served as the source for experiments in this research. This compilation includes tweets about Monkeypox from all different sources. The data characters were detokenized after the pre-processing was complete, and it was confirmed that there were no identical or null values so that the phrases could be broken up into words and labels could be given. The efficacy of the forecasting model was then assessed using CNN-LSTM models for deep learning. A comprehensive diagram of the entire system is shown in Fig. 1. The proposed framework is a three-phase process. The first phase involves data acquisition and data pre-processing. The tweets collected undergo stop-word elimination, stemming, lemmatization, and tokenization. The tokenized word then passes through the GloVe model to attain the words' glove vector, which is subsequently turned to embedding matrix of size (506,786, 300). Further the labels in the tweets are one-hot encoded. The output of embedding layer (size 506,786 $\times$ 20 $\times$ 300) of input sequence 20 and embedding dimension 300 is fed to the second phase which consists of CNN-LSTM combined framework. In this phase, the input is first fed to CNN with 20 $\times$ 30 dimensions. We then used a dropout layer of 20 $\times$ 30 to filter the CNN output before passing it to the MaxPooling layer of 5X30 dimension. The resultant is fed to LSTM of 5 X 100 size. Here the model learns the intricate features of processed tweets. The output of LSTM passes through the dropout layer of 5 $\times$ 100 dimension, processed to be fed to a fully connected layer of 4 layers of neurons with 200, 100, 50 and 30 neurons, respectively. The last phase of the process involves the classification aided by the softmax activation function.

3.1 Pre-processing

Before conducting empirical analyses on the compiled tweets, it is imperative to undertake pre-processing as a preliminary measure. The corpus of tweets gathered exhibits a lack of organization, uneven distribution, and a significant presence of commonly occurring words with little semantic value. Thus, it is imperative to

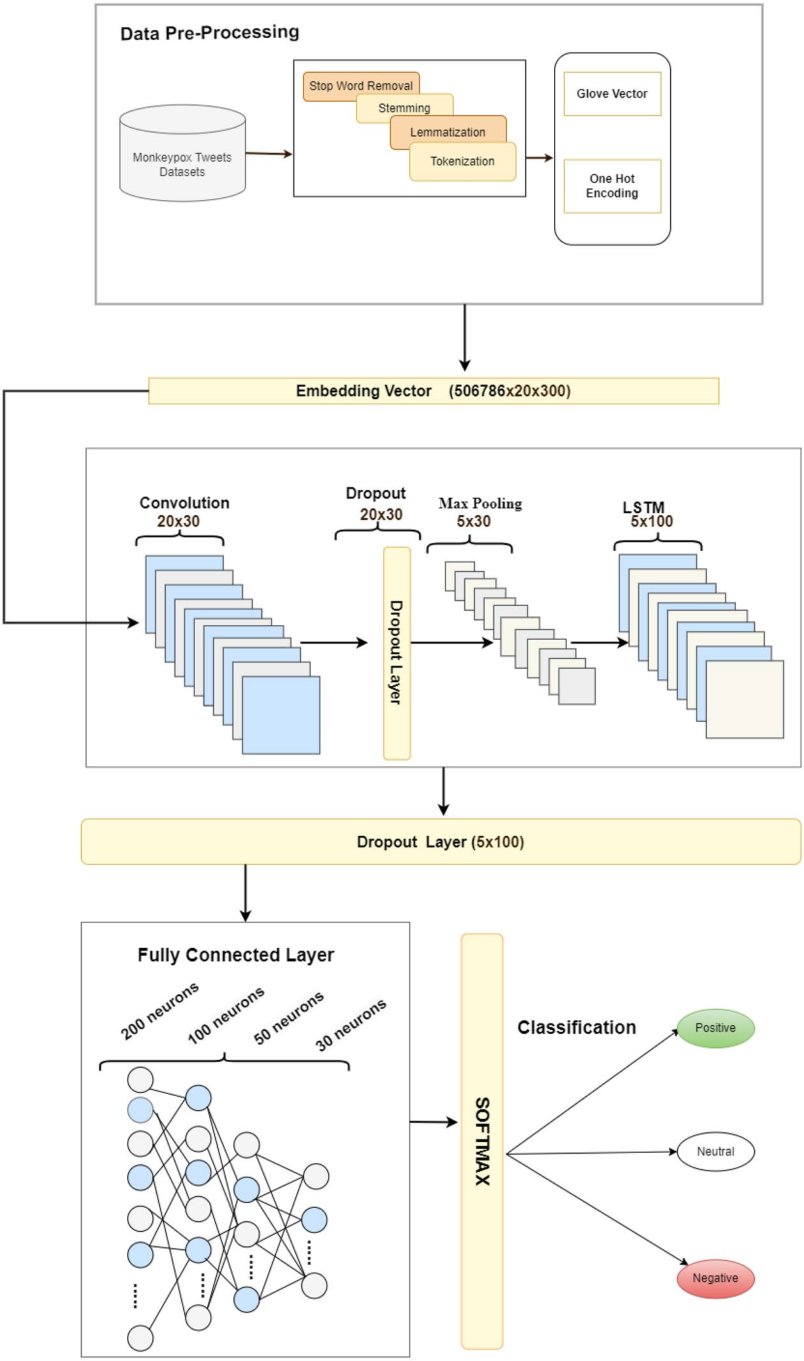


Fig. 1 CNN-LSTM architecture for sentiment analysis

pre-process all the tweets by removing any extraneous information to perform classification and prediction tasks. Since Twitter is an unstructured platform allowing multilingual publishing, data pre-processing is paramount. The present investigation was limited to the acquisition of solely English-written tweets. Pre-processing encompasses techniques such as stop word removal, word stemming, data filtering, and feature extraction [30].

The stop-word removal technique filtered the data eliminating any extraneous words. For this, we utilize a predefined list of stop words. Each tweet is checked against the available stop word list, and any terms that appear on both lists are eliminated from the respective tweets. These terms do not provide any positive contributions to the model's functioning. The resultant output is scanned for punctuation. Because the vast majority of the punctuation that may be used in tweets is meaningless, this step removes all of the punctuation symbols, including, ^, ?. Tweets might consist of URLs and hyperlinks that are removed from the next step. Further, we identify and remove hashtags, RT (retweet symbols), query phrases, and special characters. The last step involves stemming. For this Porter Stemmer [36] is used stemming from each tweet's phrases. All the terms are changed to the root word throughout this procedure. The code snippet below illustrates the approach utilized (Fig. 2).

3.2 CNN-LSTM Architecture

Assume that a string represents the word vectors and that $P_i \in R^k$ represents the K-dimensional vector comparable to the ith token in a user review of total size n,

```
# adding wouldn't type of words into stopwords list
for s in stop_words:
 new_stop_words.add(s.replace('\'',''))
 pass
stop_words=new_stop_words


# removing special symbols
base_filters='\n\t!"#$%&()*+,-./:;⟺?[\]^_`{|}~ '
word_sequences=[]
for i in x:
 i=str(i)
 i=i.replace('\'', '')
 newlist = [x for x in text_to_word_sequence(i,filters=base_filters, lower=True) if not x.startswith("@")]
 filtered_sentence = [w for w in newlist if not w in stop_words]
 word_sequences.append(filtered_sentence)
 pass


#Tokenizing words to word indices
tokenizer = Tokenizer()
tokenizer.fit_on_texts(word_sequences)
word_indices = tokenizer.texts_to_sequences(word_sequences)
word_index = tokenizer.word_index

#padding word_indices
 x_data=pad_sequences(word_indices,maxlen=MAX_SEQUENCE_LENGTH)
```

Fig. 2 Tweet pre-processing code snippet where *n* is the number of tokens in Eq. (1). Zero padding is added if the sentence is smaller than *n* characters.

$$\mathrm{P}_{1:\,\mathrm{n}} = \mathrm{P}_1 + \mathrm{P}_2 + \mathrm{P}_3 \ldots + \mathrm{P}_\mathrm{n} \tag{1}$$

The + operator denotes a concatenation operation in Eq. (1). Similarly, suppose the concatenation of the words P_i, $\mathrm{P}_{\mathrm{i+1}}$, $\mathrm{P}_{\mathrm{i+2}}$,...$\mathrm{P}_{\mathrm{i+j}}$ is equivalent to $\mathrm{P}_{\mathrm{i:\,i+j}}$. Let $\mathrm{W} \in \mathrm{R}^{\mathrm{hk}}$ denote the convolutional filters applied in an nxK dimensional matrix of a sentence with a window or gap of h words to produce a new feature matrix. The basic element $\mathrm{P}_{\mathrm{i:\,i+j}}$ denotes the local feature matrix from the ith to the (i + j)th line of the present sentence vector. Equation (2) can produce a feature C_{if} from a window of words $\mathrm{P}_{\mathrm{i:\,i+h-1}}$.

$$\mathrm{C}_\mathrm{i} = \mathrm{f}\left(\mathrm{W} \cdot \mathrm{P}_{\mathrm{i\,:\,i+h-1}} + \mathrm{b}\right) \tag{2}$$

An activation function is employed at the point where the bias, represented by b, is a real number. Examples of activation functions include hyperbolic tangent and sigmoid. The set of numbers in question contains the element "b". To produce a feature map using Eq. (3), it is necessary to apply the convolved convolutional filter to each window that contains words.

$$C = \left[C_1, C_2, C_3, C_{n-h+1}\right] \tag{3}$$

where C belongs to $\mathrm{R}^{\mathrm{n-h+1.}}$

The abovementioned Equation depicts constructing a solitary feature map utilizing a singular convolutional filter. Similarly, a convolutional layer with multiple *m* filters will generate *m(n–h + 1)* features. To effectively capture prolonged dependencies, the characteristics are promptly fed into the LSTM layer before the fully connected layer. The proposed framework combined usage of CNN and LSTM for analyzing the sentiments from the tweet's dataset. CNN has proved its mantle in varied applications involving NLP. But it suffers certain limitations in handling the above dataset. First, tweets are sequential, each containing a single data point. CNN can capture only the local patterns in data but not the long-term dependencies and sequential information. While LSTM specializes in capturing such dependencies and hence aid the CNN model. Second, we require a fixed-size input for the CNN model to excel. Tweets vary in length, ranging from a few words to a maximum limit of 288 characters. Modeling such data could result in losing vital information from the longer tweets. LSTM, unlike CNN, can process variable input lengths, making this CNN-LSTM combination an indispensable option.

Third, the data from Twitter often consists of slang, misspellings, and other non-conventional language semantics. CNN might not be able to capture the nuances of such data, while LSTM can selectively process important words and phrases inherent in tweets. Fourth, our work relies on successfully capturing the tone and context of the tweets, a task LSTM is best suited for. Thus, the proposed combination of CNN-LSTM for sentiment analysis enhances the ability of CNN and provides better results. The summary of the hybrid CNN-LSTM model is shown in Fig. 3.

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 20, 300)	152035800
conv1d_2 (Conv1D)	(None, 20, 30)	9030
dropout_11 (Dropout)	(None, 20, 30)	0
lstm_2 (LSTM)	(None, 20, 50)	16200
dropout_12 (Dropout)	(None, 20, 50)	0
flatten_2 (Flatten)	(None, 1000)	0
dense_9 (Dense)	(None, 100)	100100
dropout_13 (Dropout)	(None, 100)	0
dense_10 (Dense)	(None, 50)	5050
dropout_14 (Dropout)	(None, 50)	0
dense_11 (Dense)	(None, 30)	1530
dropout_15 (Dropout)	(None, 30)	0
dense_12 (Dense)	(None, 3)	93

Total params: 152,167,803
Trainable params: 132,003
Non-trainable params: 152,035,800

Fig. 3 Summary of the hybrid CNN-LSTM model

4 Experiment Results and Discussion

4.1 Dataset Description

According to this specific research, there is an association between Monkeypox and tweets. The dataset comprises 556,427 tweets about Monkeypox that were disseminated on Twitter on 7 May 2022 to 9 October 2022 [25]. The dataset comprises diverse tweets authored by individuals, which convey their positive, negative, and neutral sentiments. Figure 4 illustrates the classification of tweets according to their respective categories. The classification scheme depicted in the illustration assigns individuals holding neutral opinions to the 0 class, while individuals holding negative opinions are assigned to the -1 class, and those holding positive opinions are assigned to the 1 class.

4.2 Experimental Setting

Python is utilized on a computing device operating on X86-64 Ubuntu 18.04.4 LTS. The central processing unit (CPU) in use is an Intel(R) Core (TM) i7-8550U, running at a clock speed of 1.80 GHz, and the system has a total of 16 gigabytes of

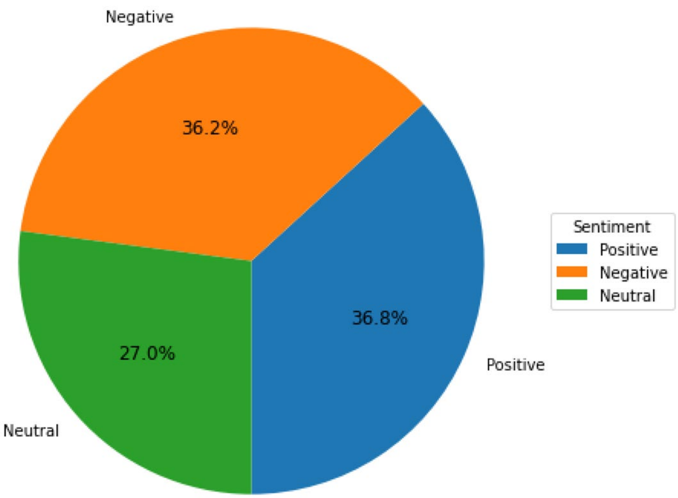


Fig. 4 Classes distribution

RAM. The proposed CNN-LSTM model is employed in this configuration to examine affective states associated with the sentiment. The CNN architecture underwent training to identify its distinguishing features. Subsequently, additional features of the LSTM framework were employed to categorize individuals' emotions.

4.3 Performance Measure

Proper evaluation procedures are necessary to mitigate the bias inherent in the differentiation of models. Precision (P), recall (R), F1 score, accuracy, and AUC are some of the metrics that are used for the classification standard the most commonly [26, 27]. The acronym (P) denotes precision, while (R) represents recall. The accuracy of the results is determined by the ratio of correctly identified samples to the total number of evaluated samples. The recall metric refers to the proportion of correctly identified positive samples from the dataset's total number of positive samples. The figure above has the potential to be further subdivided into discrete samples. The F1 score is computed through the harmonic mean of the recall and accuracy scores. One method for determining accuracy involves calculating the ratio of correctly identified samples to the total number of samples [28]. The AUC is a statistical measure that can be employed to characterize the Receiver Operating Characteristic (ROC) curve [29]. The performance of a classifier in distinguishing multiple data groups is evaluated using a measurement. The mathematical expressions of different measures are presented in Table 1.

The confusion matrix is a tabular representation utilized to evaluate the efficacy of a classification algorithm. The utilization of a confusion matrix enables the representation and concise summarization of the efficacy of a classification algorithm (Table 2). In Table 2, TN represents the count of accurately identified negative samples, whereas TP represents the count of accurately identified positive samples.

Table 1 Evaluation expressions

Measure	Expression
$\boldsymbol{P}$	$\frac{TP}{TP+FP}$
$\boldsymbol{R}$	$\frac{TP}{TP+FN}$
$\boldsymbol{Accuracy}$	$\frac{TP+TN}{TP+TN+FP+FN}$
$\boldsymbol{F1-Measure}$	$\frac{2*P*R}{P+R}$
$\boldsymbol{AUC}$	$\frac{\left(R-\frac{FP}{FP+TN}+1\right)}{2}$

Table 2 Confusion matrix

	Actual value	
	Positive	Negative
Predicted Value		
Positive	TP	FP
Negative	FN	TN

Within the field of statistics, the term "FP" refers to the count of instances in which unfavorable events were inaccurately identified as positive. Conversely, an FN-predicted value is one in which the outcome is negative, although the actual value is positive. Table 2 presents a comprehensive enumeration of the parameters of the confusion matrix.

4.4 Findings

To assess the efficacy of the hybrid CNN-LSTM architecture, we conducted a comparative analysis of the experimental results of those obtained from diverse machine learning models. The findings enabled us to assess the efficacy of the proposed model. In the comparison process, the evaluation metrics considered include accuracy, AUC, f1-score, precision, recall, and confusion matrix. The proposed model currently under discussion employs a range of classifications encompassing neutral, positive, and negative categories. The accuracy and loss history of various deep learning-based models' test sets is demonstrated in Figs. 5, 6, and 7 after each training epoch for the monkeypox tweet dataset.

The results indicate that the hybrid CNN-LSTM architecture performs better than simpler CNN and LSTM-based deep learning models. At various epochs, the accuracy of the hybrid CNN-LSTM model was evaluated. Figure 7 shows the outcomes of the same. The outcome demonstrates that accuracy grew along with the number of epochs. The accuracy of the suggested hybrid model is 85% after 20 epochs and 91% after 140 epochs.

The accuracy depicted in Fig. 7 exhibits a steady increase starting from epoch 20 and maintains a consistent level until epoch 120, approximately 91%. The outcomes are evaluated against CNN and LSTM models. These two models' accuracy varies on epochs. The loss versus epoch variation for several models is shown in Figs. 5, 6, and 7. Figure 7 shows that the loss of the models considerably decreased as the

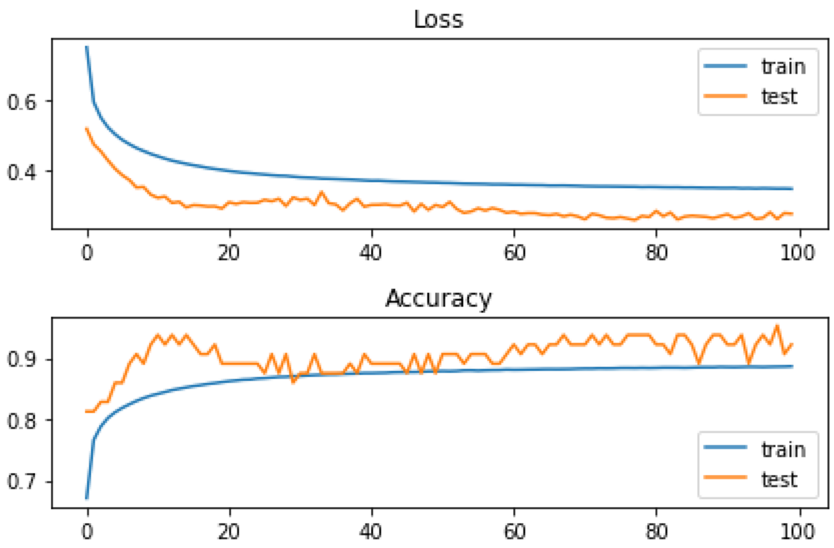


Fig. 5 The *loss* and *accuracy* of the CNN model

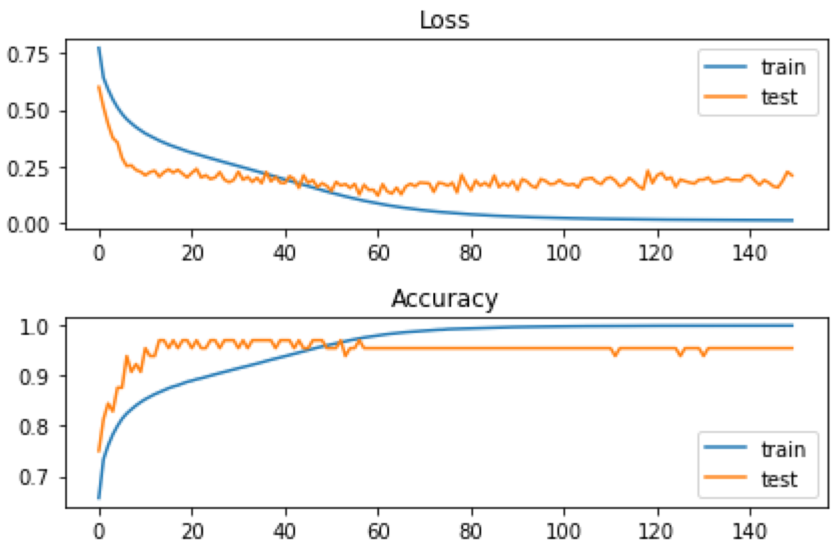


Fig. 6 The *loss* and *accuracy* of the LSTM model

number of epochs increased. Figures 5 and 6 also show the disappearance of CNN and LSTM. These models lose more money than the hybrid model does.

The most favorable outcome of the experiment will be deemed acceptable. Eight models' findings from the monkeypox tweets dataset are displayed in Table 3 and Fig. 8. In terms of Accuracy, Precision, Recall, F1-score, and AUC across all five assessment parameters, the suggested technique performs better than existing

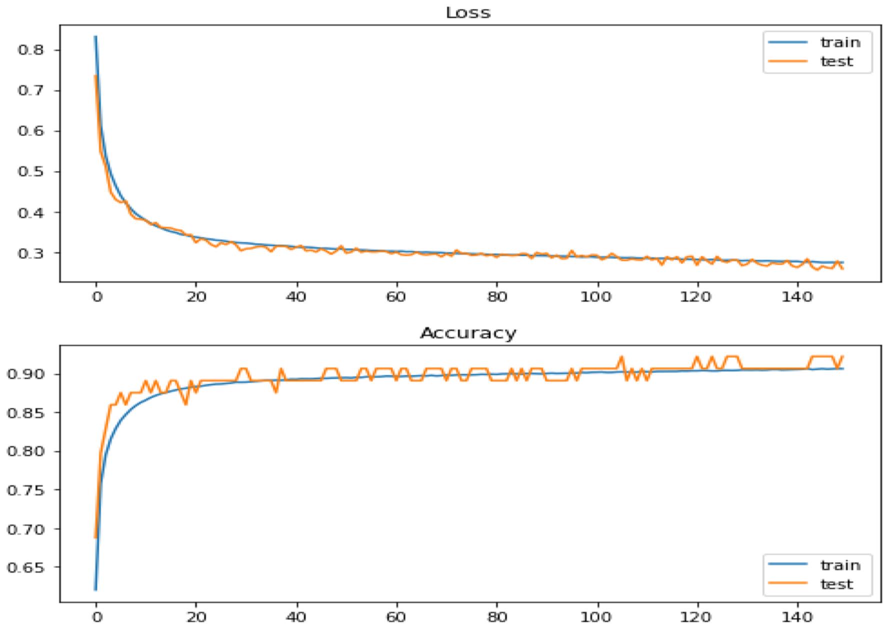


Fig. 7 The *loss* and *accuracy* of the hybrid CNN-LSTM

Table 3 Performance comparison of different models

Model	Accuracy	Precision	Recall
CNN-LSTM	0.91	0.91	0.91
CNN	0.87	0.89	0.87
LSTM	0.86	0.86	0.86
Naive Bayes	0.66	0.77	0.64
SVM	0.73	0.75	0.71
Logistic Regression	0.90	0.89	0.89
Random Forest	0.69	0.75	0.66
Decision Tree	0.86	0.79	0.80

state-of-the-art methods overall. After epoch 60, the loss for the CNN-LSTM is roughly 0.1, and accuracy can approach 91%. This shows that our suggested model has more accuracy, indicating that it can accurately assess different sentiments.

Table 3 compares the various models' accuracy, precision, and recall performance. The hybrid CNN-LSTM model performs better than deep learning and machine learning models. Using the fi-score and AUC measures, we have also contrasted our findings. Figure 8 illustrates this outcome. It concludes that the suggested model outperforms other state-of-the-art models. It obtains a 91.24% fi-score and a 92% AUC, whereas the logistic regression model only achieves an 89% accuracy and a 91% AUC.

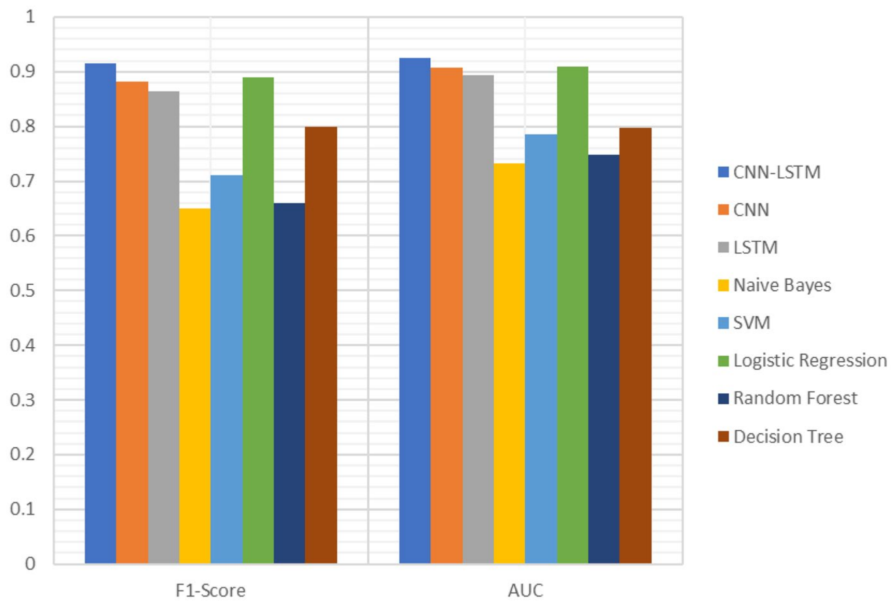
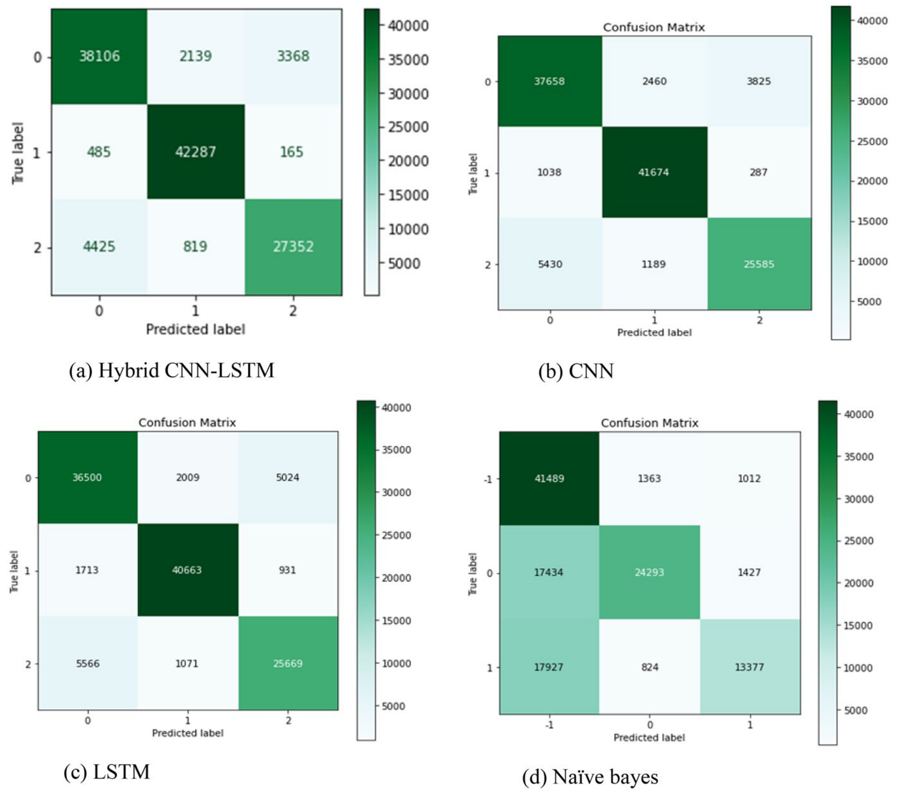


Fig. 8 F1-score and AUC comparison

Figures 9 and 10 display the confusion matrices generated on the hidden test dataset to help the reader better understand the classification performance of each model. The CNN-LSTM model has the highest classification accuracy, as seen in Fig. 9a. From the comparative analysis hybrid CNN-LSTM model classifies 38,106 samples accurately in the positive class, 42,287 samples in the negative class, and 27,352 in the neutral class as opposed to all other models. The misclassifications for positive class to negative are 2139 and to neutral are 3368, respectively. Further, 485 samples in the negative class were misclassified to be in a positive class, while 165 negative class samples were misclassified as neutral class samples. 4425 neutral class samples were misclassified in the positive class, and 819 neutral class samples were misclassified in the negative class, which is quite less than other models; hence the proposed model is superior to others.

5 Discussion

The outcomes generated by the hybrid CNN-LSTM model that was constructed are juxtaposed with the results yielded by several other models that are widely acknowledged as state-of-the-art. Furthermore, this compilation comprises models that employ machine-learning techniques. The research results suggest that the CNN-LSTM model generates superior quality outcomes compared to alternative models. The process of tokenization of the data is crucial as it strips the data into small parts suitable to be embedded into vector space. For better training, the batch size is set to 64, and the epochs to achieve high accuracy are set to 140. Upon application to the



(a) Hybrid CNN-LSTM

(b) CNN

(c) LSTM

(d) Naïve bayes

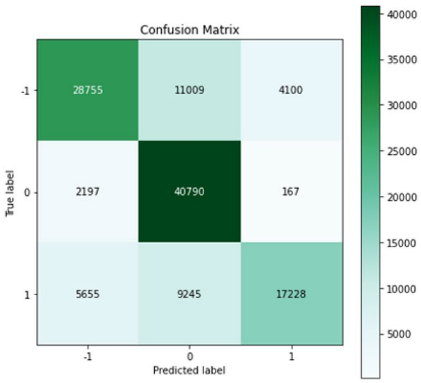
Fig. 9 The obtained confusion matrix of different models

dataset of tweets about Monkeypox, the LSTM model generates an f1-score measure that is accurate 91% of the time.

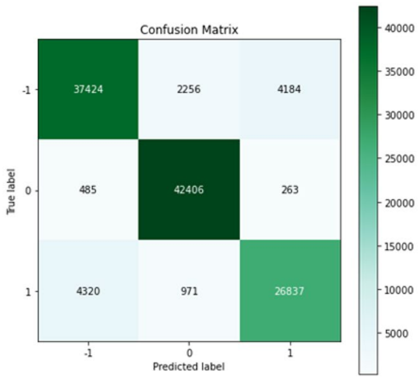
Furthermore, enhancements have been made to both the precision and recall of the suggested model, resulting in increased accuracy. Comparing several correct classifications, the confusion matrix shows the model's strength. Hybrid CNN-LSTM models have achieved better accuracy than traditional machine learning models. The CNN layers can extract relevant features from the text embeddings, while LSTM layers can capture the temporal sequence of tweets to learn how the language used in the tweets changes over time.

6 Conclusion and Future Work

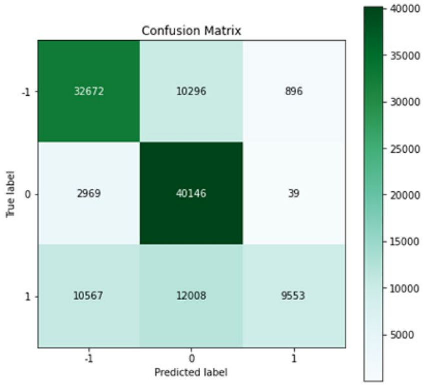
Our study's results show how CNN and LSTM methods can be used to assess tweets and identify the emotional polarity of their content. The findings of this research are based on whether participants view the contagious disease monkeypox as good, negative, or neutral. We were able to determine how exactly and



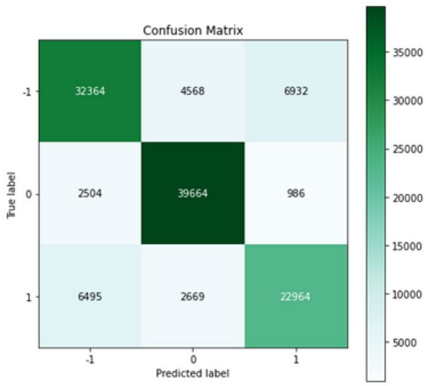
(a) SVM



(b) Logistic Regression



(c) Random forest



(d) Decision tree

Fig. 10 The obtained confusion matrix of different models

effectively we could predict and evaluate user opinion on social media platforms like Twitter. When compared to the currently employed techniques, the hybrid CNN-LSTM performed better. The combined CNN-LSTM design with hyperparameter tunings obtained a degree of accuracy of 91% for the datasets related to Monkeypox. The results show that the suggested method is adequate and appropriate for classifying the emotions in tweets about Monkeypox. The proposed effort also benefits society by increasing knowledge of the newly discovered monkeypox virus. The features of a system may be optimized using an iterative approach that makes use of the more reliable methods, which may be improved even further in succeeding studies. It is possible to achieve this optimization. It is also possible to alter it so that topic identification and mood categorization can be made simultaneously. It would aid in the system's development in the future.

Funding Not applicable.

Data availability The data supporting this study's findings are available on request from the corresponding author.

Declarations

Conflict of Interest The authors of this manuscript state that they have no conflict of interest.

Ethical approval Not Applicable.

References

1. Golbeck, J., Robles, C., Edmondson, M., Turner, K.: Predicting personality from twitter. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. IEEE; 2011, p. 149–156.
2. Quercia, D., Kosinski, M., Stillwell, D., Crowcroft, J.: Our twitter profiles, our selves: Predicting personality with twitter. In: 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing. IEEE; 2011, p. 180–185
3. Lokesh, S., Kumar, P.M., Devi, M.R., Parthasarathy, P., Gokulnath, C.: An automatic tamilspeech recognition system by using bidirectional recurrent neural network with self-organizing map. Neural Comput. Appl. **31**(5), 1521–1531 (2019)
4. Chauhan, S., Banerjee, R., Chakraborty, C., Mittal, M., Shiva, A., Ravi, V.: A self-congruence and impulse buying effect on user's shopping behaviour over social networking sites: an empirical study. Int. J. Pervasive Comput. Commun. (2021). https://doi.org/10.1108/IJPCC-01-2021-0013
5. Mondal, A., Mahata, S., Dey, M., Das, D.: Classification of COVID19 tweets using machine learning approaches. In: Proceedings of the Sixth SocialMediaMining for Health (# SMM4H) Workshop and Shared Task. Mexico City, 2021; p. 135–7.
6. Ashok Kumar P., Shankar, G.S., Gautham, S., Reddy, M.P.K., Reddy, G.T.: A two-stage text feature selection algorithm for improving text classification. In: ACM Trans Asian Low-Resour Lang Inf Process. New York, NY: Association for Computing Machinery, 2021;. p. 19. https://doi.org/10.1145/3425781.
7. Hakak, S., Alazab, M., Khan, S., Gadekallu, T.R., Maddikunta, P.K.R., Khan, W.Z.: An ensemble machine learning approach through effective feature extraction to classify fake news. Future Gener. Comput Syst. **117**, 47–58 (2021). https://doi.org/10.1016/j.future.2020.11.022
8. Yang, L., Zhang, H., Li, D., Xiao, F., Yang, S.: Facial expression recognition based on transfer learning and SVM. J. Phys. Conf. Ser. **2025**(1), 012015 (2021)
9. https://www.who.int/news-room/fact-sheets/detail/monkeypox. Accessed 12 Jun 2022
10. https://openwho.org/courses/monkeypox-introduction. Accessed 12 Jun 2022
11. WHO Calls Emergency Meeting as Monkeypox Cases Top 100 in Europe Available online: https://www.reuters.com/world/europe/monkeypox-outbreak-europe-largest-ever-region-cases-cross-100-2022-05-20/. Accessed 12 June 2022
12. Unlikely Monkeypox Outbreak Will Lead to Pandemic, WHO Says Available online: https://www.reuters.com/business/healthcare-pharmaceuticals/unlikely-monkeypox-outbreak-will-lead-pandemic-says-who-2022-05-30/. Accessed 12 June 2022
13. Kelleher, S.R.: CDC Raises Monkeypox Travel Alert to Level 2 Available online: https://www.forbes.com/sites/suzannerowankelleher/2022/06/07/cdc-raises-monkeypox-travel-alert-to-level-2/?sh=269eee1e3f93. Accessed 12 June 2022
14. Kuvvetli, Y., Deveci, M., Paksoy, T., Garg, H.: A predictive analytics model for COVID-19 pandemic using artificial neural networks. Decis. Anal. J. **1**, 100007 (2021)
15. Wang, J., Wang, M.: Review of the emotional feature extraction and classification using EEG signals. Cognit. Robot. **1**, 29–40 (2021)
16. Cambria, E.: Affective computing and sentiment analysis. IEEE Intell. Syst. **31**(2), 102–107 (2016)
17. Liu, X.-Q., Wu, Q.-L., Pan, W.-T.: Sentiment classification of micro-blog comments based on randomforest algorithm. Concurr. Comput. Pract. Exp. **31**(10), e4746 (2019)

18. Hassan, A.; Mahmood, A.: Deep Learning approach for sentiment analysis of short texts. In: Proceedings of the 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), Nagoya, Japan, 24–26 April 2017; pp. 705–710.
19. Shen, Q.; Wang, Z.; Sun, Y. Sentiment Analysis of Movie Reviews Based on CNN-BLSTM. In: International Conference on Intelligence Science; Springer: Berlin/Heidelberg, Germany, 2017; pp. 164–171.
20. Meena, G., Mohbey, K.K., Indian, A.: Categorizing sentiment polarities in social networks data using convolutional neural network. SN Comput. Sci. **3**(2), 116 (2022)
21. Chhajer, P., Shah, M., Kshirsagar, A.: The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. Decis. Anal. J. **2**, 100015 (2022)
22. Praveen, S.V., Ittamalla, R., Deepak, G.: Analyzing Indian general public's perspective on anxiety, stress and trauma during Covid-19—a machine learning study of 840,000 tweets. Diabetes Metab. Syndr. Clin. **15**, 667–671 (2021). https://doi.org/10.1016/j.dsx.2021.03.016
23. Lyons, M. J., Kamachi, M., Gyoba, J.: Coding facial expressions with Gabor wavelets (IVC special issue). arXiv preprint arXiv:2009.05938. (2020)
24. Lyons, M. J.: "Excavating AI" Re-excavated: Debunking a Fallacious Account of the JAFFE Dataset. arXiv preprint arXiv:2107.13998. (2021)
25. Thakur, N.: MonkeyPox2022Tweets: a large-scale Twitter dataset on the 2022 Monkeypox Outbreak, findings from analysis of tweets, and open research questions. Infect. Dis. Rep. **14**(6), 855–883 (2022)
26. Mohbey, K. K., Sharma, S., Kumar, S., & Sharma, M.: COVID-19 identification and analysis using CT scan images: deep transfer learning-based approach. In: Sudeep Tanwar (ed.) Blockchain Applications for Healthcare Informatics, pp. 447–470. Academic Press (2022)
27. Xie, Y., Xing, F., Kong, X., Su, H., & Yang, L.: Beyond classification: structured regression for robust cell detection using convolutional neural network. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 358–365. Springer, Cham (2015)
28. Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.
29. Mohbey, K.K.: Multi-class approach for user behavior prediction using deep learning framework on twitter election dataset. J. Data Inform. Manag. **2**(1), 1–14 (2020)
30. Li, M., Ch'ng, E., Chong, A.Y.L., See, S.: Multi-class Twitter sentiment classification with emojis. Ind. Manag. Data Syst. **118**(9), 1804–1820 (2018)
31. Malik, S., Jain, S.: Knowledge-infused text classification for the biomedical domain. Int. J. Inform. Syst. Model. Des. (IJISMD) **13**(10), 1–15 (2022)
32. Meena, G., Mohbey, K.K., Kumar, S., Lokesh, K.: A hybrid deep learning approach for detecting sentiment polarities and knowledge graph representation on monkeypox tweets. Decis. Anal. J. **7**, 100243 (2023)
33. Gruenwald, L., Jain, S., Groppe, S. (eds.): Leveraging Artificial Intelligence in Global Epidemics. Academic Press (2021)
34. Dash, S., Chakravarty, S., Mohanty, S.N., Pattanaik, C.R., Jain, S.: A deep learning method to forecast COVID-19 outbreak. New Gener. Comput. **39**(3–4), 515–539 (2021)
35. Hura, G.S., Groppe, S., Jain, S., Gruenwald, L.: Artificial intelligence in global epidemics, part 2. New Gener. Comput. **40**, 935–939 (2022)
36. Jahanbin, K., Jokar, M., Rahmanian, V.: Using twitter and web news mining to predict the monkeypox outbreak. Asian Pac. J. Trop. Med. **15**(5), 236 (2022)