ORIGINAL ARTICLE

TSA-CNN-AOA: Twitter sentiment analysis using CNN optimized via arithmetic optimization algorithm

Serpil Aslan[1] · Soner Kızıloluk[2] · Eser Sert[2]

Abstract

COVID-19, a novel virus from the coronavirus family, broke out in Wuhan city of China and spread all over the world, killing more than 5.5 million people. The speed of spreading is still critical as an infectious disease, and it causes more and more deaths each passing day. COVID-19 pandemic has resulted in many different psychological effects on people's mental states, such as anxiety, fear, and similar complex feelings. Millions of people worldwide have shared their opinions on COVID-19 on several social media websites, particularly on Twitter. Therefore, it is likely to minimize the negative psychological impact of the disease on society by obtaining individuals' views on COVID-19 from social media platforms, making deductions from their statements, and identifying negative statements about the disease. In this respect, Twitter sentiment analysis (TSA), a recently popular research topic, is used to perform data analysis on social media platforms such as Twitter and reach certain conclusions. The present study, too, proposes TSA using convolutional neural network optimized via arithmetic optimization algorithm (TSA-CNN-AOA) approach. Firstly, using a designed API, 173,638 tweets about COVID-19 were extracted from Twitter between July 25, 2020, and August 30, 2020 to create a database. Later, significant information was extracted from this database using FastText Skip-gram. The proposed approach benefits from a designed convolutional neural network (CNN) model as a feature extractor. Thanks to arithmetic optimization algorithm (AOA), a feature selection process was also applied to the features obtained from CNN. Later, K-nearest neighbors (KNN), support vector machine, and decision tree were used to classify tweets as positive, negative, and neutral. In order to measure the TSA performance of the proposed method, it was compared with different approaches. The results demonstrated that TSA-CNN-AOA (KNN) achieved the highest tweet classification performance with an accuracy rate of 95.098. It is evident from the experimental studies that the proposed approach displayed a much higher TSA performance compared to other similar approaches in the existing literature.

Keywords COVID-19 · Convolutional neural network · Arithmetic optimization algorithm · Sentiment analysis · Twitter

✉ Serpil Aslan
serpil.aslan@ozal.edu.tr

Soner Kızıloluk
soner.kiziloluk@ozal.edu.tr

Eser Sert
eser.sert@ozal.edu.tr

[1] Department of Software Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, 44210 Malatya, Turkey

[2] Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, 44210 Malatya, Turkey

1 Introduction

In the late 2019, increase in number of pneumonia cases in Wuhan city of China was reported to World Health Organization (WHO). In January 2020, the disease was officially named as COVID-19. On March 11, 2020, COVID-19 was accepted as a global pandemic by WHO. As of January 23, 2020, more than 346 million confirmed cases and 5.5 million deaths have been reported around the world [1]. COVID-19 is spread through respiratory droplets in various situations such as sneezing, coughing and speaking. The virus can live up to a few days on plastic surfaces and up to a few hours on cardboard surfaces [2].

The emergence of its symptoms may take 2 to 14 days in an infected individual [3]. The most common symptoms of COVID-19 are dry cough, fever or fatigue [2]. In addition to these, some less common symptoms of the disease are body aches, diarrhea, sore throat and headache [4].

Taking and obeying various vital measures against COVID-19 such as vaccination, wearing masks, social distancing, and personal hygiene have played a huge role in controlling the pandemic. Twitter is one of the foremost social media platforms which help raise awareness of these crucial issues. Individual feelings against the pandemic bear utmost importance in implementing strategies that will eliminate the pandemic [5]. In line with this, the analysis of social media platforms offers valuable information for health staff and government decision-makers in a country. Social media analysis also contributes to the identification of massive emotional changes in society and the elimination of a potential social crisis by raising awareness of current social problems [6, 7]. When the studies in the current literature are analyzed, it can be observed that social media analysis is a popular research topic and that tweets have been usually classified as positive, negative, or neutral in these studies [8, 9]. CNN has also been recently used in Twitter sentiment analysis (TSA) and displayed remarkable success in the classification of tweets. The present study, similarly, focuses on performing TSA on Twitter users' tweets on COVID-19; thanks to a CNN-based approach. The proposed TSA-CNN-AOA optimizes hyperparameters of CNN via arithmetic optimization algorithm (AOA) in order to increase the classification performance of CNN. Thus, TSA-CNN-AOA approach was used to classify tweets on COVID-19 as positive, negative, and neutral.

As the number of Twitter users has been increasing substantially over the past few years, large masses can be informed about recent news in a very short period of time, making it possible to find out their sensitivity toward a certain issue. In parallel with this, sentiment analysis (SA) on Twitter data through various machine learning (ML) techniques has been a popular research trend in recent times. Various researchers have developed so far different ML classification approaches for TSA [10]. Since Twitter hashtags can be effectively used to identify and distinguish between different topics on the platform, they play a certain role in performing TSA on a given topic. From a historical perspective, it can be stated that TSA has been employed using different approaches. For instance, in 2009, Go et al. [11] combined the Naive Bayes classifier and n-gram language model in order to divide tweets into three different groups as positive, neutral, and negative. Pak and Paroubek [12] performed the automatic collection of a corpus for sentiment analysis and opinion mining to propose a multinomial Naive Bayes-based sentiment classifier using N-gram and POS-tags, which divided tweets into three groups objective, positive and negative.

In 2011, Kouloumpis et al. [13] explored the utility of micro-blogging and lexicon features in a three-way sentiment classifier. Similarly, Xia et al. [14] proposed an ensemble framework for sentiment classification and designed two different schemes of feature sets, i.e., "POS-based feature sets" and "WR-based feature set." In this study, Naive Bayes (NB), maximum entropy (ME), or support vector machines (SVM) were used as component classification models in the ensemble system. In this way, they applied three different ensemble methods, i.e., fixed combination, weighted combination and meta-classifier combination, to sentiment classification [14].

Pagolu et al. [15] benefited from two different feature extract techniques, namely Word2vec and N-gram for TSA, and applied SA and ML approaches to tweets to analyze the relationship between movements of a company on the stock market and related sentiments in various tweets. In recent times, developments in the field of deep learning (DL) have paved the way for their use in CNN [16] and the LSTM variant of recurrent neural networks [17, 18].

The main contributions of the present study are summarized as follows:

(1) The present study proposes a novel approach for TSA of people's thoughts about the COVID-19 pandemic on Twitter, which is one of the most important agendas today.

(2) Within the framework of the present study, an API was designed to extract 173,638 tweets about COVID-19 from Twitter between July 25 and August 30, 2020. Due to their unsuitability for data processing as raw data, they were subjected to preprocessing to omit numerous special characters, statements, links, emoji, etc., which could otherwise affect experimental studies negatively during the analysis.

(3) In order to analyze the effect of COVID-19 pandemic on the society in a more detailed manner, tweets obtained from Twitter under four different hashtags ("#covid19, #coronavirus #pandemic and #covid19vaccine") were divided into two different topics as pandemic (#covid19, #coronavirus #pandemic) and covid19vaccine. Later, sentiment distribution in each topic was analyzed on their own. The analysis results suggest that the effect of pandemic on the society significantly affected people's sentimental tendencies toward COVID-19 vaccination process.

(4) Significant information was extracted from this database using FastText Skip-gram model.

(5) The present study proposes Twitter sentiment analysis using convolutional neural network optimized via arithmetic optimization algorithm (TSA-CNN-AOA). The proposed approach relies on the CNN model as a feature extractor. After a feature selection process is applied to high-level local features obtained from CNN using arithmetic optimization algorithm (AOA), one of the most recent meta-heuristic optimization algorithms, K-nearest neighbors (KNN), SVM, and decision tree, was used to classify tweets as positive, negative and neutral.

(6) The proposed approach identifies negative sentiments on the social media platform, i.e., Twitter, and tags inappropriate, missing, or incorrect information about COVID-19, thus reducing the possibility of misinformation to a minimum level. The results of the present study will broaden individuals' general views on the different vaccines and the pandemic itself.

The rest of the present study is organized as follows: Sect. 2 presents a literature review on the related works in this field. Section 3 describes the obtained data set, pre-processing steps applied to this data set and related theoretical frameworks. Section 4 defines the proposed approach and analyzes the experimental results of the proposed approach on the given data set comparatively. Section 5 concludes the study.

2 Related work

2.1 Sentiment analysis

SA is a combination of data mining and text mining as two different research fields and aims to find out sentiments expressed in a written language [19]. It mainly focuses on the automatic extraction of subjective information conveyed through a certain text [20] and performs many different tasks such as sentiment extraction, sentiment classification, sentiment summarization, and so forth. SA, which is also known as opinion mining, can be categorized into three different levels: (1) document level [21], (2) sentence level [22], and (3) aspect-based level [23]. At the document level, SA attempts to identify sentiment polarities in a given text. The most critical point here is the assumption that the document in question focuses on a single topic or entity. At the sentence level, SA tries to determine whether a sentence expresses a positive, negative or a neutral opinion. Additionally, it also defines sentences with subjective or objective sentiment. Finally, at the aspect-based level, SA fulfills three main tasks: entity/object identification, feature extraction, and feature polarity.

SA is a fairly challenging task, as human language involves many different factors such as countless grammatical variations, idiomatic expressions, slang use, misspelling, synonymous and ambiguous words, all of which make it arduous to analyze it in detail. For instance, it is usually difficult for a SA model to analyze synonymous words in different contextual settings or words with different semantic aspects. Therefore, stemming techniques are often used to overcome these challenges, as they help analyzers find the root of a given the word. Even though SA succeeds in eliminating many different linguistic problems, it may not truly analyze an opinion when a different word is used, which may decrease its overall performance.

Current SA methods can be divided into three categories [24]: ML-based, dictionary-based, and DL-based SA. ML-based approaches usually utilize a bag of words to convert texts to features [14]. Later, features obtained from complex ML approaches are fed into classifiers such as Naive Bayes (NB), decisions trees (DT), and support vector machine (SVM) [25]. Dictionary-based approaches usually collect positive and negative sentiment words in a given text to calculate text polarity based on the sum of these words [26]. Unlike dictionary-based approaches, ML-based approaches may benefit from sentiment dictionaries, which consist of a range of positive and negative values assigned to different words [24]. In this respect, ML-based approaches offer various advantages compared to dictionary-based approaches. In the existing literature, hybrid ML- and dictionary-based methods have been used together [24]. ML-based approaches were later replaced by DL-based approaches, whose experimental results seem to be more promising when compared to other approaches [27, 28].

DL-based approaches have been popular among many researchers due to their considerable success in SA. For example, Chen et al. [29] proposed a single-dimensional CNN model in which temporal relations were embedded into user and product representations to improve SA performance at the document level. Similarly, Liu et al. [30] proposed an artificial neural network-based approach that recommends idioms in essay writing. This model calculates similarities between the given context and candidate idioms. Klachbrenner proposed a CNN model to analyze introductory sentences at varying lengths [28]. Tai et al. proposed long short-term memory (LTSM) with feedback features by improving RNN architecture [31]. Likewise, Schuster and Paliwal proposed a Bi-LSTM model based on two different LSTM networks by improving these networks [32].

Kumar et al. [33] proposed Plutchik wheel of emotion-based approach for textual SA using word embedding and Plutchik's wheel of emotions. Villavicencio et al. [34] proposed a SA approach to classify sentiments in positive, neutral, and negative polarities toward COVID-19 vaccines in Philippines. In this study, the obtained data were pre-processed using several NLP approaches to develop a SA classification model through Naive Bayes classification approach using RapidMiner data science software and thus help the government take decisions related to the vaccination schedule. Shamrat et al. [35] observed that people express their opinions on the reliability and effectiveness of COVID-19 vaccines in social media platforms such as Twitter and extracted such tweets from the website using Twitter API authentication token. Later, following a data processing step using NLP, they classified the processed data using a supervised KNN classification algorithm and divided them into three different categories as positive, negative and neutral. Sontayasara et al. [36] proposed a SVM-based TSA approach.

Twitter is one of the most popular social media websites in which people express and share their feelings about a specific topic. However, the available data on Twitter are usually too big and unstructured to handle, which makes it often demanding to analyze and extract subjective information from them. In recent years, SA has been employed in a number of fields such as business, politics, and social media. Similar to many other social media platforms, SA is one of the leading tools for gaining insight into different individuals' opinions and views on various topics. It helps companies and government agencies collect information on people's opinions and decisions in an easier way. Using DL-based natural language processing (NLP) techniques, the present study also relies on the vastness and availability of Twitter data in order to analyze Twitter users' sentiments about COVID-19 through their tweets and comments on these tweets.

2.2 Deep learning models for sentiment analysis

In the current literature, DL-based approaches have been widely applied to various text data for SA recently. Ankita et al. [37] proposed a CNN-LSTM model to perform SA on #BlackLivesMatter tweets in two different US states and divided them into two different categories as hateful and non-hateful, which yielded a classification accuracy rate of 94%. Usama et al. [38] proposed a new model architecture based on RNN with CNN-based attention for SA using three different datasets and achieved an accuracy rate of 83.64%, 51.14%, and 89.62% on these datasets. Behera et al. [39] proposed a hybrid model combining CNN and LSTM for SA on customer reviews, which displayed an

accuracy rate of 94.90% in four different customer review datasets.

Khasanah [40] proposed a single-layered CNN model with FastText embedding for SA on text data. This model was tested using the Model Movie Review (MR) dataset and the Stanford Sentiment Treebank (SST2), which yielded a classification accuracy rate of 80% and 83.9%, respectively. Jain et al. [41] proposed CNN-LSTM to perform classification in Airlinequality Airline Sentiment Data and Twitter Airline Sentiment Data and achieved an accuracy rate of 87.6% and 87.5%, respectively. Onan [42] combined TF-IDF weighted Glove word embedding with CNN-LSTM architecture and used the proposed model for SA on product reviews obtained from Twitter to divide them into two categories positive and negative. Jain et al. [43] proposed a hybrid bidirectional long short-term memory and a softmax attention layer and convolution neural network (softAttBiLSTM-feature-richCNN) for sarcasm detection. This model was tested on political and entertainment content in Twitter data, reaching an accuracy rate of 92.71%.

Nezhad and Deihimi [44] created two different datasets from tweets in the Persian language about the COVID-19 vaccine developed by Iran (COVIran Barekat) and imported vaccines (AstraZeneca/Oxford, Pfizer/BioNTech, Moderna, Sinopharm, etc.) between April 1, 2021, and September 30, 2021. Afterward, they used a DL-SA model based on CNN-LSTM architecture to extract tweets and categorize them as positive, negative, and neutral to reveal monthly changes in sentiment from a statistical perspective. Behl et al. [45] proposed a multilayer perceptron (MLP) network model to identify basic humanitarian needs in tweets during emergency situations and natural disasters. Their proposed model was trained using three different datasets, namely 'resource needs,' 'resource availability,' and 'others,' obtained from tweets about Nepal Earthquake and Italy Earthquake. The trained model was later tested on a different tweet dataset about COVID-19 and displayed a classification accuracy rate of 83%.

Basiri et al. [46] proposed a novel method based on the fusion of four DL and one conventional supervised machine learning model for SA on tweets about COVID-19 in eight different countries and reported statistical and temporal changes in sentiments, as well as sentiment differences in these eight countries. Sitaula et al. [47] brought three different fastText-based, domain-specific, and domain-agnostic-based CNN models together to propose a new model for SA on tweets about COVID-19 in Nepal. AlBadani et al. [48] proposed a SA approach using deep learning models by combining the Universal Language Model Fine-Tuning (ULMFiT) and SVM to detect people's attitudes based on their comments. Vernikou et al. [49] focused on TSA for the classification of user sentiments in

tweets about COVID-19 on Twitter and implemented SA using seven different deep learning models based on LSTM neural networks.

In the present study, a CNN architecture was designed as a feature extractor. However, unlike studies mentioned in the literature review above, AOA, which is one of the most recent meta-heuristic algorithms, was used for feature extraction for the data obtained from CNN architecture. Afterward, SVM, decision tree, and KNN methods were used for the classification process.

3 Preliminaries

3.1 FastText word embedding vector

FastText is a Word2Vec-based model developed by Facebook for text classification in 2016. It transforms words or texts into continuous vectors which can be used in any language (spoken language) for a given task. The main difference between this method and Word2Vec is that it splits words into a few character-based "n-grams" instead of giving them as single inputs to the artificial neural network. Thus, it may achieve a semantic similarity that cannot be achieved by Word2Vec [50]. Similar to Word2Vec, FastText offers two different models: Skipgram and CBOW. While the skip-gram model uses neighboring words to predict a target word, CBOW relies on all context words in order to predict a target word [51]. Both methods create a text file that contains numerical representations (i.e., vectors) of learned words. The present study benefits from a FastText Skip-gram model. A 300-dimensional vector space and a sub character dimension $n = 3$ were used in FastText configuration.

3.2 TextBlob

TextBlob is a Python library used for many different NLP tasks such as parts of speech tagging, SA, noun phrase extraction, translation, and classification [52, 53]. TextBlob library returns two main features of a sentence "Polarity" and "Subjectivity" [54]. In sentence subjectivity, while an objective sentence expresses phenomenal information about the world, a subjective sentence expresses personal sentiments and beliefs. Subjective sentences usually reflect various personal sentiments such as beliefs, wishes, opinions, doubts, delights, or fears. Subjectivity is a variable parameter between a range of [0, 1]. A subjectivity value converges to 0 points to a more factual and objective sentence, whereas a higher subjectivity value makes it an opinion. Sentence polarity can be defined as positive, negative, or neutral sentimental orientation in written or verbal language. Polarity is assigned a value between a

range of $[-1, 1]$. -1, 0, and 1 represent negative, neutral, and positive statements, respectively.

3.3 Convolutional neural network (CNN)

CNN is different from artificial neural networks in that it is a DL approach consisting of different layers with feature extraction. Figure 1 shows a typical CNN architecture. It is widely used for the classification of image contents. It is based on artificial neural networks and possesses a customized DL architecture. Images are given as inputs to CNN architecture. Similar to artificial neural network, CNN was also inspired by the working principles of human brain. As shown in Fig. 1, CNN architecture contains convolution layers, pooling layers, a fully connected layer, and an activation layer. The first layer of a CNN architecture is a convolutional layer that extracts local features from an image. Due to its architectural system, the pooling layer in the nth location is connected to a fully connected layer. There are a few backpropagation steps in CNN during the learning process to minimize losses. As shown in Fig. 1, some activation functions such as Softmax and Tanh are also used to obtain the output.

3.4 The arithmetic optimization algorithm

Proposed by Abualigah et al. [55], AOA was inspired by four basic mathematical operations: addition (A), subtraction (S), multiplication (M), and division (D). The main steps of AOA are described briefly in the following sections.

3.4.1 Initialization process

Similar to other population-based meta-heuristic algorithms in the existing literature, AOA uses an initial population consisting of candidate solutions with random values. In each iteration, the candidate solution with the best fitness value in the population is called the best obtained solution. The initial population (X) is represented as a matrix, as can be seen in Eq. 1:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,n} \\ x_{2,1} & x_{2,2} & \ldots & x_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{N,1} & x_{N,2} & \ldots & x_{N,n} \end{bmatrix} \quad (1)$$

where N is the number of candidate solutions, whereas n is the problem dimension.

Later, in order to select a search phase (exploration or exploitation) function for the algorithm, the operator called Math Optimizer Accelerated (MOA) is calculated according to Eq. 2:
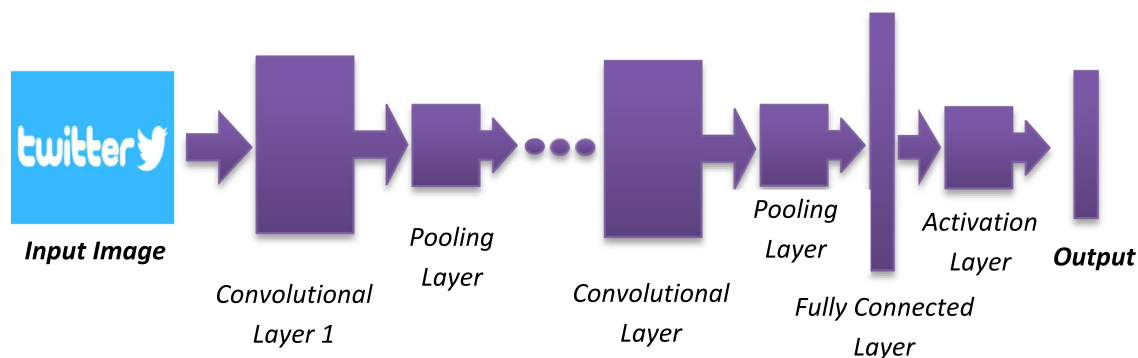
Fig. 1 CNN architecture

$$\text{MOA}(C\_\text{iter}) = \text{Min} + C\_\text{iter} \times \left(\frac{\text{Max} - \text{Min}}{M\_\text{iter}}\right) \quad (2)$$

Here, while C_iter represents the current iteration, M_iter denotes the maximum number of iteration, and, finally, Min and Max are the minimum and maximum value of the accelerated function, respectively [55].

3.4.2 Exploration and exploitation phases

The selection of exploration or exploitation phase of AOA depends on MOA value in Eq. 2. After a random *r1* value is created between 0 and 1, exploration is selected if $r1 > \text{MOA}$, while exploration is selected if $r1 < \text{MOA}$. Operator M and D are used in the exploration phase of AOA to facilitate search in a wide range of search space. The exploration phase of the AOA is given in Eq. 3 [55]:

$$\begin{aligned} &x_{i,j}(C\_\text{iter}+1) \\ &= \begin{cases} \text{best}(x_j) \div (\text{MOP}+\varepsilon) \times \left((\text{UB}_j - \text{LB}_j) \times \mu + \text{LB}_j\right), & \text{if } r2 < 0.5 \\ \text{best}(x_j) \times \text{MOP} \times \left((\text{UB}_j - \text{LB}_j) \times \mu + \text{LB}_j\right), & \text{else} \end{cases} \end{aligned}$$

$$(3)$$

Here, $x_{i,j}(C\_\text{iter} + 1)$ is the candidate solution in the following iteration, while $\text{best}(x_j)$ is the jth dimension of the current best candidate. UB_j and LB_j represent the upper and lower bound values of the jth dimension, respectively. *r2* is a random value between 0 and 1, ε is a very small positive number, and, finally, μ is the control parameter with a value of 0.5 and used for the exploration phase. Operator S and A are used in the exploitation phase of AOA in order to facilitate local search surrounding the best candidate solution. The exploitation phase of the AOA is given in Eq. 4 [55]:

$$\begin{aligned} &x_{i,j}(C\_\text{iter}+1) \\ &= \begin{cases} \text{best}(x_j) - \text{MOP} \times \left((\text{UB}_j - \text{LB}_j) \times \mu + \text{LB}_j\right), & \text{if } r3 < 0.5 \\ \text{best}(x_j) + \text{MOP} \times \left((\text{UB}_j - \text{LB}_j) \times \mu + \text{LB}_j\right), & \text{else} \end{cases} \end{aligned}$$

$$(4)$$

Here, *r3* is a random value between 0 and 1. Math

Optimizer probability (MOP) in Eqs. 3 and 4 is a coefficient and calculated as given in Eq. 5:

$$\text{MOP}(C\_\text{iter}) = 1 - \frac{C\_\text{iter}^{1/\alpha}}{M\_\text{iter}^{1/\alpha}} \quad (5)$$

Here, C_iter is the current iteration, whereas M_iter is the maximum number of iteration. α is a sensitive parameter with a value of 5. Flowchart of the AOA is shown in Fig. 2. [55]. Firstly, an initial population is created with AOA parameters (M_iter, μ, etc.) and candidate solutions with random values. Later, the fitness, MAO and MOP values of candidate solutions are updated in each iteration. MOA helps select between exploration and exploitation phases. During the exploration phase, operator D or M randomly is applied via a randomly generated r2 parameter. During the exploitation phase, on the other hand, operator S or A randomly is applied via a randomly generated r3 parameter. The algorithm is ended once it reaches a maximum number of iterations, and the candidate solution with the best fitness value is accepted as the solution.

Maintaining the appropriate balance between exploration and exploitation strategies is crucial for a metaheuristic optimization algorithm to achieve high performance. Exploration is the ability to explore the search space globally. On the other hand, exploitation is the ability to find better solutions by performing a local search in the immediate vicinity of a candidate solution. According to Abualigah et al. [55], an appropriate balance between exploration and exploitation strategies was achieved by applying exploration strategies with M and D operators and exploitation strategies with S and A operators in the AOA recommended by [55]. Also, Abualigah et al. [55] tested AOA on 29 benchmark functions and five engineering design problems in an experimental study. They compared the test results with the following algorithms [55]: genetic algorithm (GA), particle swarm optimization (PSO), biogeography-based optimization (BBO), flower pollination algorithm (FPA), gray wolf optimizer (GWO), bat algorithm (BAT), firefly algorithm (FA), cuckoo search
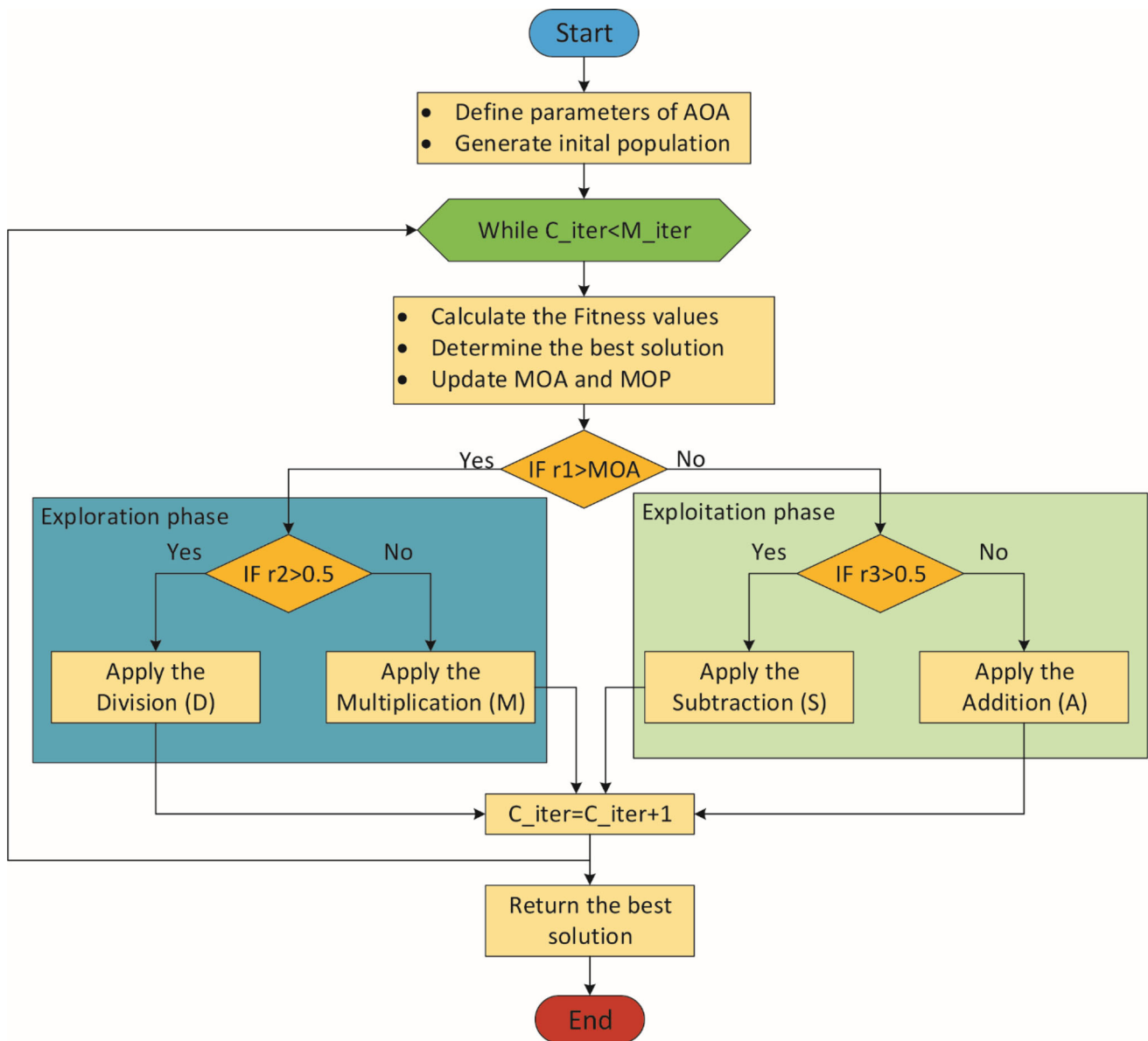
Fig. 2 Flowchart of the AOA

algorithm (CS), moth-flame optimization (MFO), gravitational search algorithm (GSA), differential evolution (DE).

Abualigah et al. [55], as a result of these test process, proved that AOA was more successful than these 11 compared algorithms. Therefore, AOA was preferred in this study.

3.5 Machine learning-based approaches

ML studies the ways in which computers can be trained by a dataset. ML approaches are often used in text mining studies to fulfill the task of classifying different types of texts through scientific features. A wide range of ML classifiers can be used to perform SA, which is one of the

popular topics in text mining. Many different ML classifiers are available in Python Scikit-Learn Library [56]. Because it is an open-access library, it appeals to a wide range of users worldwide. The following classifiers used in the present study to compare the performance of the proposed approach with other approaches were obtained from Python Scikit-Learn Library: SVM, Naive Bayes, KNN, decision tree, and logistic regression (LR).

3.5.1 Support vector machine (SVM)

Proposed by Cortes and Vapnik, SVM is a binary classification tool that can be extended for multiclass problems [57]. In the present study, it was used for a tripartite

classification as positive, negative, and neutral. SVM is a powerful technique used for various problems such as nonlinear classification, regression, and detection of outliers. However, it aims to cross-validate the data and display a poor performance in small-size datasets.

The SVM can divide data into two or more classes with linear separation mechanisms in two-dimensional space, planar separation mechanisms in three-dimensional space, and hyperplane separation mechanisms in multidimensional space. The method, which is frequently used in the determination of linearly separable classes, is useful in the classification of nonlinear data by moving the input space that cannot be parsed linearly to higher-dimensional linearly separable space; thanks to kernel functions. SVM mainly aims to maximize the distance between support vectors from various classes that are separated by a boundary line known as a hyperplane. Support vectors are instances of classes that are most closely related to the hyperplane. They define the class to which they belong by lying on a plane parallel to the hyperplane.

3.5.2 Naive Bayes

Proposed by Thomas Bayes, Naive Bayes is a controlled ML learning classifier that calculates the probability for a feature through statistical methods [58]. It is employed for many different purposes, such as diagnostic classification, text and document classification, spam e-mail listing, and classification and prediction-based models. Naive Bayes classifiers aim to predict features based on the assumption that they are not interdependent.

Naive Bayes is one of the simplest, understandable and easily applicable machine learning algorithms used in text classification, which is created using Bayes' theorem. Using this method, it is possible to find the probability that a sample belongs to the class value of the target attribute.

The formula for the Naive Bayes algorithm is as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$P(A|B)$ represents the probability of event A that occurs when event B occurs, whereas $P(B|A)$ represents the probability that event B will occur if event A occurs. Therefore, $P(A)$ and $P(B)$ are a priori probabilities of events A and B.

3.5.3 K-nearest neighbors (KNN)

Proposed by Cover and Hart, K-nearest neighbors (KNN) is a sample-based and controlled ML algorithm [59]. Because it possesses the capacity to support multi-label classification, it is widely preferred in SA studies. KNN calculates

the distance between a target value and other values to find K-nearest neighbor values [60]. It requires two main parameters for functioning: K value and distance metric. It has always been a popular algorithm due to its simplicity and classification success.

An example of KNN classification algorithm is shown in Fig. 3. The red dot in the figure is classified by green squares or blue triangles. When $K = 3$ in Fig. 3, the situation inside the circle is checked. Because the number of blue triangles will be higher than green squares, the classification is performed according to blue triangles. However, if $K = 7$, the situation in the dashed circle is taken into account, and, since the number of blue triangles will be lower than green squares, the classification is performed according to green squares.

3.5.4 Logistic regression

Logistic rgression is a multivariate statistical ML learning algorithm that classifies data by taking dichotomous outcome variables into account to create a logarithmic line that separates them from each other. As its name may cause some misunderstandings, it is often used for regression. However, as a statistical model, it can also be used for classification tasks. LR aims to establish a model that can describe the relationship between dependent and independent variables using a small number of variables. Estimating probabilities with a logistic function, LR analyzes the connection between a categorical dependent variable and one or more independent factor(s). Being one of the simplest ML algorithms, logistic regression offers high efficiency and a low variance level and can also be used for feature extraction.
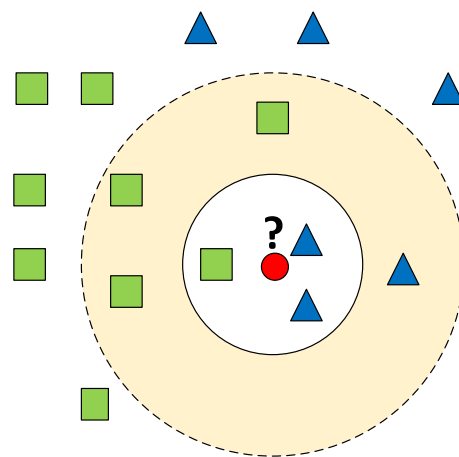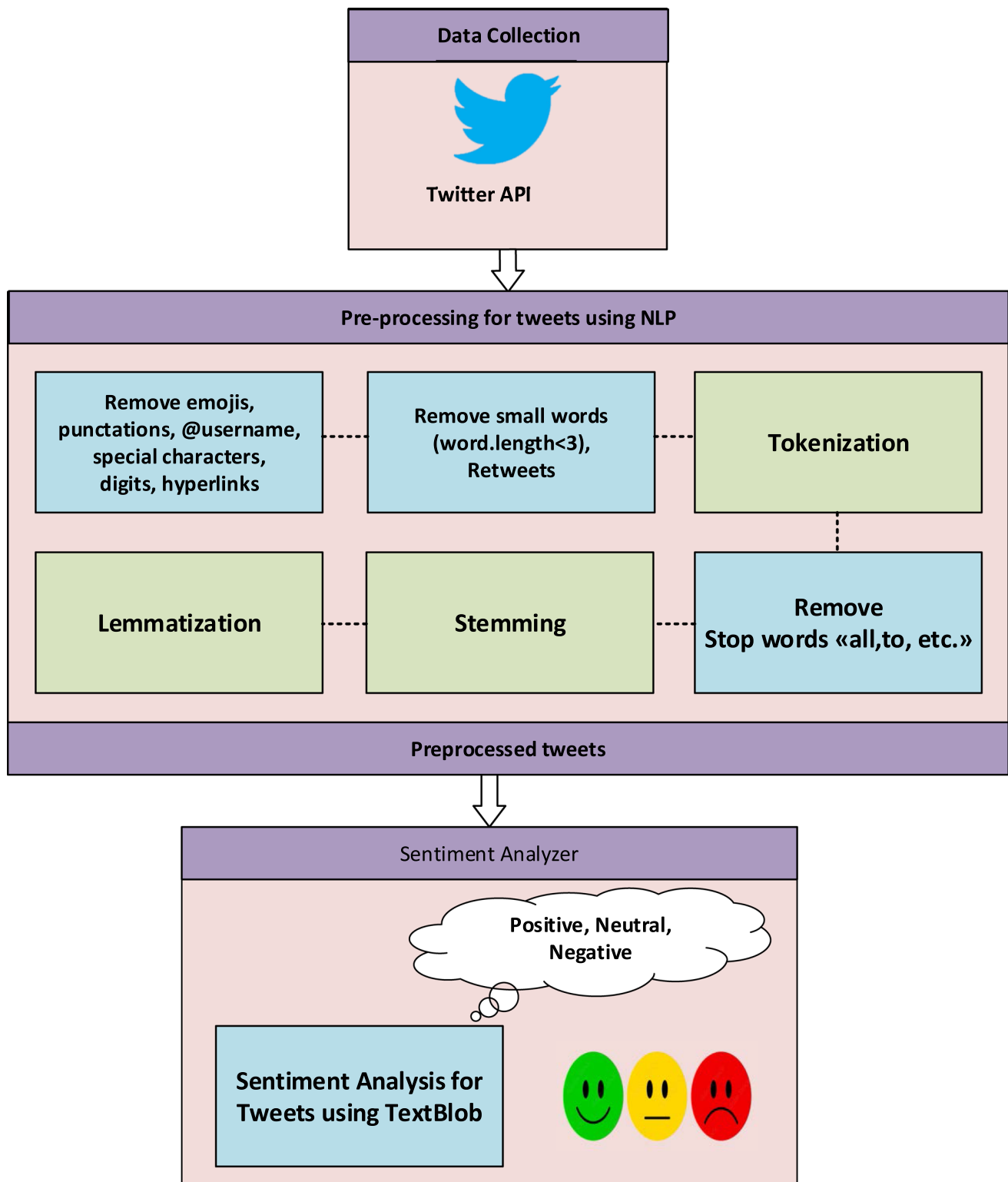

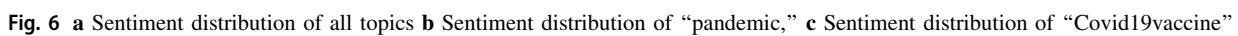**Fig. 3** An example of KNN classification algorithm

Fig. 4 Data cleansing, preprocessing and sentiment analyzer

3.5.5 Decision tree

Decision tree is a classification algorithm that consists of decision and leaf nodes and creates a tree-like classification model. It splits very large datasets into smaller subsets based on specific decision rules. The existing literature often refers to four different forms of DT algorithms: ID3, C4.5, C5.0, and CART. The ID3 algorithm, one of the

Fig. 5 **a** The word cloud for raw tweets **b** The word cloud for the preprocessed tweets

Fig. 6 **a** Sentiment distribution of all topics **b** Sentiment distribution of "pandemic," **c** Sentiment distribution of "Covid19vaccine"

univariate decision trees, benefits from information gain approach that provides the most considerable information gain from the targets in each node. In ID3 algorithm, after trees reach their maximum size, pruning is applied to improve their ability to generalize invisible data. C4.5 algorithm eliminates the shortcomings of ID3 algorithm using the gain ratio approach, which is calculated via division of information and information gain. The C5.0 algorithm is a more effective and memory-saving variant of C4.5 method. In CART algorithm, a statistical approach, binary trees are constructed using features and thresholds that provide the most important information gained from each decision node.

It usually benefits from the CART (Classification and Regression Trees) algorithm to perform classification using decision trees. Starting with a root node, a decision tree algorithm splits each node into two leaves to form a series of binary branching nodes. Meanwhile, as each leaf connected to a node represents a class label, it is necessary to check the number of positive and negative class labels on the leaves to select the branch representing the best decision. If a single positive or negative class label is found on the leaf, the decision tree does not require splitting anymore, and top-down induction is completed [61].

4 The proposed approach

The flowchart of the proposed approach for the analysis of the effect of COVID-19 pandemic on the society through tweets is shown in Fig. 8. The approach consists of three basic steps.

Step 1: In this step, open-access tweets about COVID-19 are extracted from Twitter. Later, raw data are preprocessed to create a cleansed dataset for sentiment analysis and polarity detection in each tweet. The flowchart of the data collection and preprocessing architecture is shown in Fig. 4.

Step 2: This is the word representation step in which each tweet obtained from Step 1 is vectorized using FastText word embedding approach. Later, each word representation is converted to an image of 19 × 300. Here, 19 denotes the maximum tweet length, while 300 denotes a 300-dimensional feature matrix for each word representation.

Step 3: This is the CNN-AOA-based feature reduction step. Tweet images of 19 × 300 obtained from Step 2 are given to CNN model as inputs. A total of 3400 features are extracted from the second max_pooling layer of the proposed CNN model for each tweet. Later, AOA and three different classifiers (KNN, SVM, and decision tree) are used to perform feature selection on these 3400 features.

4.1 Datasets and preprocessing

The data in the present study were collected from Twitter, which was the largest social media platform with 199 million active users around the world in early 2021 [62]. Publicly shared Twitter messages in English about COVID-19 were used to create a dataset. MAXQDA, a qualitative data analysis tool [63], collected Twitter data about COVID-19, such as tweets, retweets, and mentions between July 25, 2020, and August 30, 2020. Firstly, the most popular trending topics about COVID-19 were
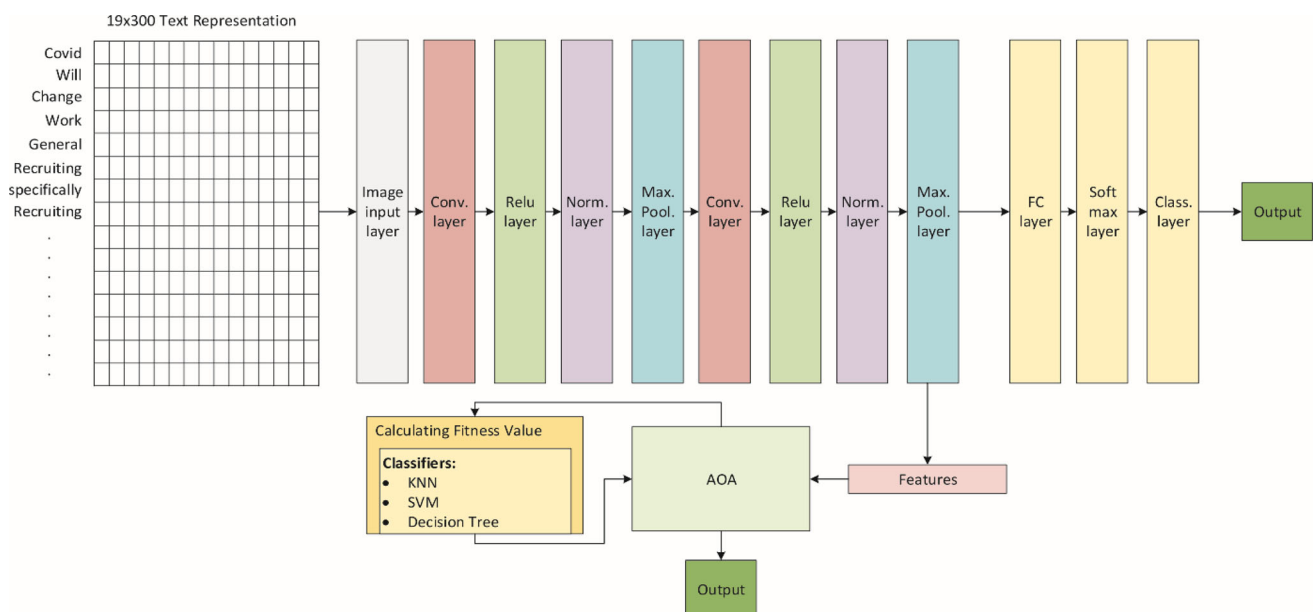

Fig. 7 CNN-AOA section of the proposed model

detected to collect related tweets. Secondly, a combination of search terms, namely "#covid19, #coronavirus #pandemic and #covid19vaccine," was used. Consequently, 173,638 tweets in English were included in the dataset. Finally, the following preprocessing was applied for data cleansing to verify the dataset.

The data obtained from Twitter, the largest social media platform, are not always cleansed. Data cleansing is a part of text mining that aims to deduce meaning from a text and omit non-analyzable words and other irrelevant components. Twitter data usually contain various irrelevant special characters, expressions, links, tags, and emoji signs which may negatively affect experimental studies during

the analysis process. In addition, such characters often pose difficulty for SA. As shown in Fig. 4, the following processes were applied to the obtained Twitter dataset for data cleansing.

Figure 5(a) shows the word cloud for the collected tweet data without any preprocessing and data cleansing. Word cloud is one of the most popular data visualization techniques to represent text data. Important and frequent textual data can be emphasized better when visualized on a word cloud. The main objective of preprocessing is to reduce the number of words in a given text without corrupting the semantic aspect of that text. It can be seen in Fig. 5(a) that the obtained raw data contain irrelevant
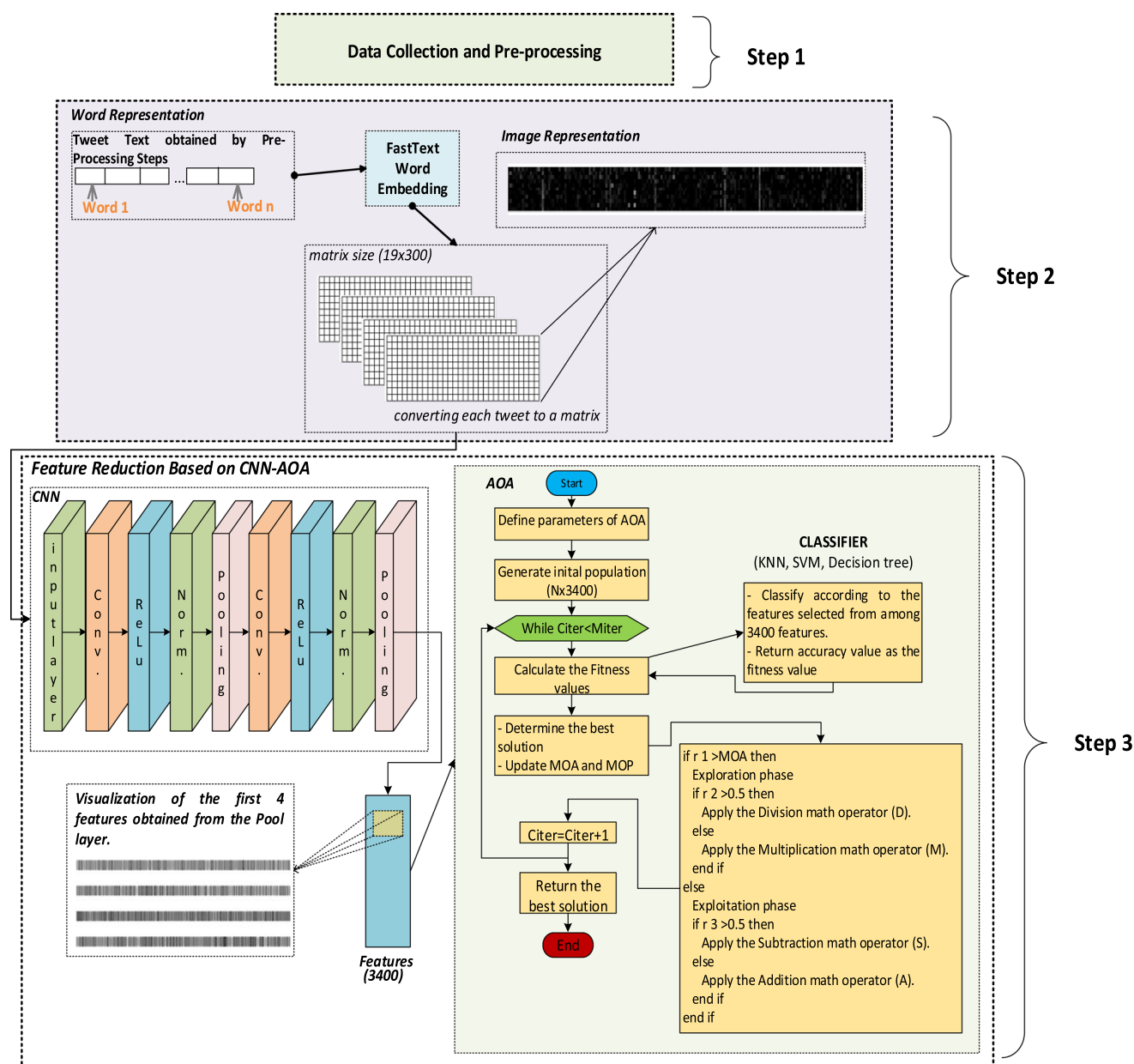
Fig. 8 The flowchart of TSA-CNN-AOA approach

Table 1 Accuracy, F1-score, precision and recall results

	Accuracy	F1-Score	Precision	Recall
TSA-CNN-AOA (Decision tree)	92.533	90.541	90.643	90.444
TSA-CNN-AOA (KNN)	95.098	93.711	94.236	93.230
TSA-CNN-AOA (SVM)	95.007	93.649	94.256	93.100
CNN	89.717	86.502	88.961	84.961
SVM	77.638	73.785	74.006	74.697
Naive Bayes	84.255	80.482	81.505	79.737
Logistic regression	88.942	85.707	87.105	84.641
Decision tree	83.464	80.108	80.086	80.135
KNN	88.434	85.227	86.677	84.121

words and phrases which do not contribute to SA process. As such, it was definitely necessary to perform preprocessing on these raw data prior to any data analysis. As shown in Fig. 5(b), the number of irrelevant and meaningless words was reduced considerably, as the number of tweets decreased from 173,638 to 147,329 following the preprocessing.

Following the preprocessing of the collected tweet dataset, the remaining 147,329 tweets were analyzed using Textblob and classified into three categories: positive, negative and neutral. The sentiment distribution based on class labels of all tweets is shown in Fig. 6(a). According to the analysis results, the number of positive, negative and neutral tweets was 54,847 (37.22%), 22,334 (15.15%) and 70,148 (47.61%), respectively. Neutral tweets represent the majority of all tweets, which can be considered as a sign of confusion and uncertainty in people's minds about COVID-19 pandemic. This is very likely to cause a negative effect on the public opinion.

For a more detailed analysis of the effect of COVID-19 pandemic on the society, tweets which were collected from Twitter on four different hashtags ("#covid19, #coronavirus #pandemic and #covid19vaccine") were categorized into two sub-topics: pandemic (#covid19, #coronavirus #pandemic) and covid19vaccine. The sentiment distribution of both sub-topics, i.e., pandemic and COVID-19 vaccine, is shown in Fig. 6(b) and (c), respectively. As shown in Fig. 6, the rate of negative sentiments is higher in vaccine sub-topic compared to pandemic sub-topic, while neutral sentiments are lower. It can be thus argued that people with a neutral sentimental attitude toward pandemic maintained a negative attitude toward vaccination. In this respect, it bears utmost importance to detect inaccurate or missing information shared on different social media platforms about the pandemic.

4.2 The proposed TSA-CNN-AOA approach

The present study proposes a TSA-CNN-AOA approach to perform SA on tweets about COVID-19. The designed CNN model was used as a feature extractor. Later, the features obtained from CNN were selected using AOA for a classification process using SVM, decision tree, and KNN methods. CNN-AOA section of the proposed model is shown in Fig. 7.

The designed CNN model benefits from word embedding as input parameters. Word embedding converts each tweet to a 19×300 matrix consisting of numerical values. Afterward, these matrices are trained using the designed CNN model, which consists of two convolution layers, two ReLUs, two cross-channel normalization layers, two max pooling layers, one fully connected layer, one softmax layer and one classification layer.

AOA is used to select among different features obtained from the second max pooling layer of the designed CNN. A total of 3400 features are obtained from the second max pooling layer and, accordingly, each candidate solution in the initial population of AOA consists of 3400 dimensions with randomly generated 0 s and 1 s. If the value of a dimension is 0 in the candidate solution, a dimension with a value of 1 is selected. For instance, randomly generated X_1 candidate solution with 3400 dimensions is represented by $X_1 = [x_{1,1} = 1, \quad x_{1,2} = 1, \quad x_{1,3} = 0, \quad x_{1,4} = 1,\ldots\ldots, x_{1,3400} = 0]$ vector. Since the value of first, second and fourth dimensions is 1 in this vector, the proposed approach will use these features. As given in Eq. 6, the initial population is represented by a matrix:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,3400} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,3400} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,3400} \end{bmatrix} \qquad (6)$$

The fitness value of each candidate solution in AOA is calculated using three different classifiers (KNN, SVM, and decision tree). The features with a value of 1 in the candidate solution are selected by the classifier, and they are used to perform training and prediction. In the end, the obtained accuracy value is accepted as the fitness value of a candidate solution. The candidate solution with the highest fitness value is given as the problem solution by the
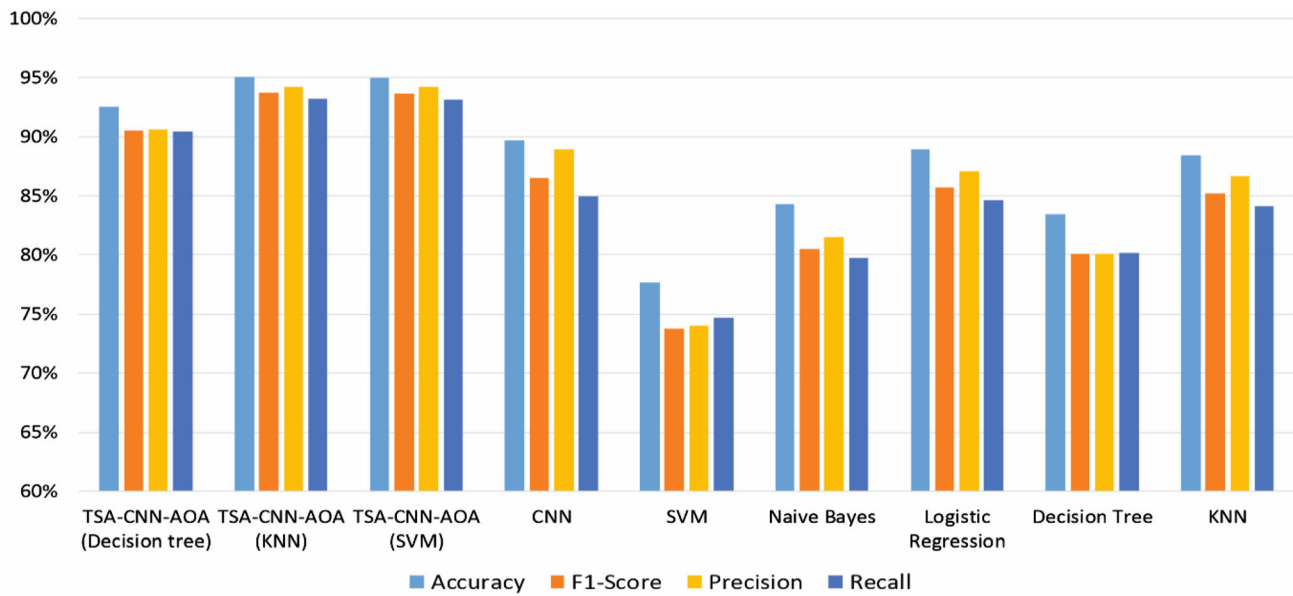
Fig. 9 Bar charts of all results

Fig. 10 Confusion matrices of all approaches

algorithm. The proposed TSA-CNN-AOA approach is shown in the flowchart in Fig. 8. Source codes of TSA-CNN-AOA are available at https://drive.google.com/drive/folders/1S3SFatKgOA0IzzITgfyNrBW7gcVwBAGx.

4.3 Experimental results

In the experimental studies, tweets about COVID-19 were classified into three groups as positive, negative, and neutral, using different methods to compare their respective
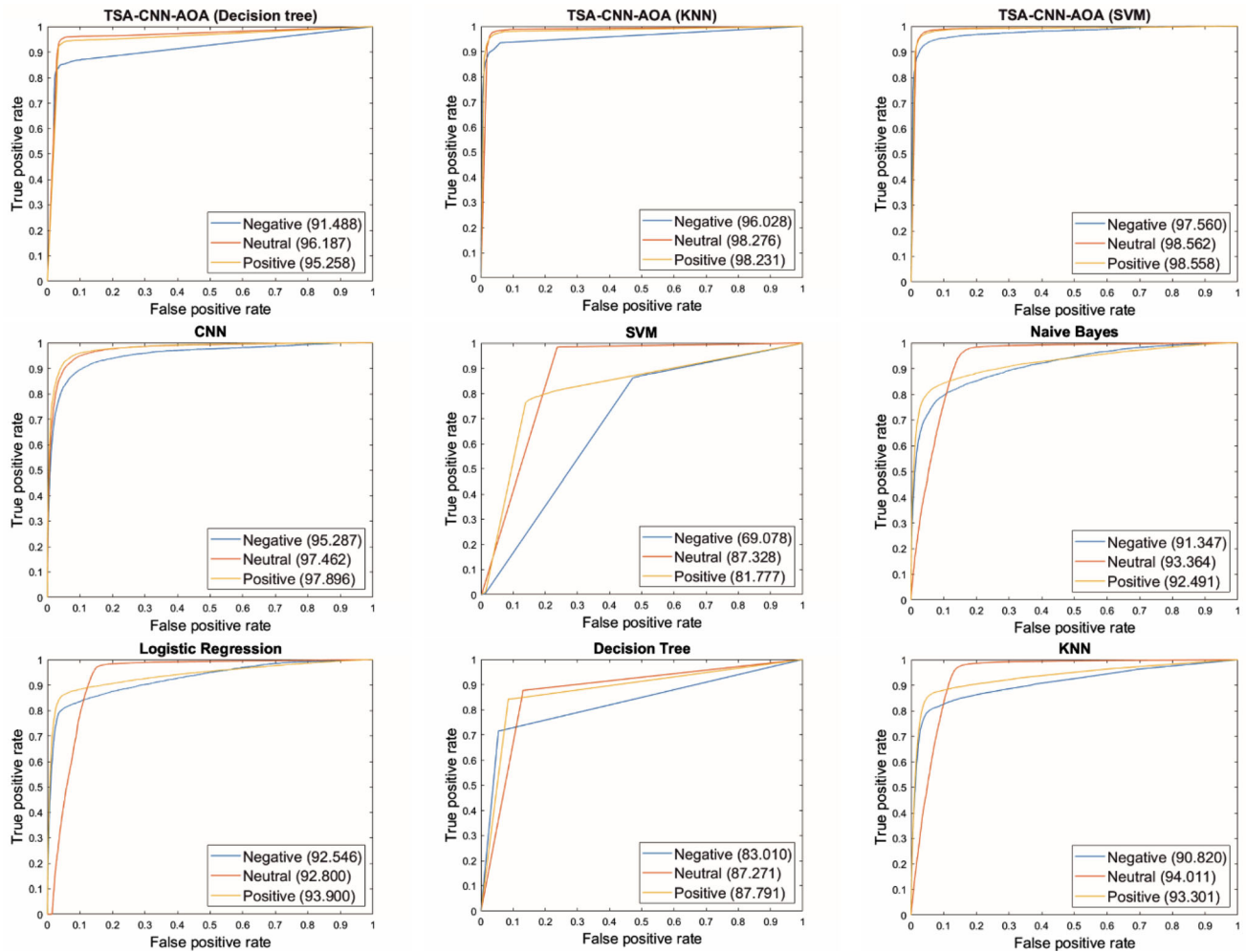
Fig. 11 ROC curves of all methods

Table 2 The accuracy rates of other proposed approaches for SA on COVID-19 in the existing literature

Model	Accuracy	Reported by
TSA-CNN-AOA (Decision tree)	0.9253	
TSA-CNN-AOA (KNN)	0.9509	
TSA-CNN-AOA (SVM)	0.9500	
SVM	0.8010	Naasem et al. [64]
NB	0.7320	Naasem et al. [64]
RF	0.8230	Naasem et al. [64]
LR	0.8300	Nair et al. [65]
Fusion model (NBSVM-CNN-BiGRU)	0.8580	Basiri et al. [46]
CNN	0.8160	Basiri et al. [46]
LSTM- PCNN	0.8399	Al-Sarem et al. [66]
SVM	0.8500	Khakharia et al. [67]
Conv1D-LSTM-Glove	0.8706	Jalil et al. [68]
KNN	0.7921	Jalil et al. [68]
RF	0.9100	Rustam et al. [69]
SVM	0.8700	Rustam et al. [69]
DT	0.9100	Rustam et al. [69]

classification performances. Firstly, the designed CNN model was used for the classification process with the following parameters: Optimizer "adam," initial learning rate "0,001," mini-batch size "128" and epoch number "1." Secondly, the designed CNN model was used as a feature extractor, and the features obtained from CNN were later used for feature selection via AOA. Thirdly, three different classifiers, i.e., KNN, SVM, and decision tree, were used to calculate the fitness value of each candidate solution in AOA, which yielded three different classification scenarios. (TSA-CNN-AOA (KNN), TSA-CNN-AOA (SVM), and TSA-CNN-AOA (Decision Tree)). The number of the initial population and maximum iterations was set to 10 for AOA. Finally, the classification process was completed using standard SVM, Naive Bayes, logistic regression, decision tree, and KNN.

Nearly 20% ($n = 32,131$) of all Twitter COVID-19 dataset were used for testing process in the present study. Accuracy, F1-Score, precision and recall performances of all approaches for the test dataset in the experimental study are given in Table 1. The results are also presented in a detailed bar graph in Fig. 9. It can be observed that the highest classification accuracy was achieved by TSA-CNN-AOA (KNN), followed by TSA-CNN-AOA (SVM) with an accuracy rate of 95.007%. On the other hand, the classification accuracy rates of CNN, Naive Bayes, logistic regression, decision tree and KNN ranged between 83 and 89%. The lowest performance belongs to SVM with 77%.

The experimental results demonstrated that feature selection via AOA following feature extraction via CNN significantly contributes to the classification performance. While the classification accuracy of CNN was 89.717%, the accuracy rates increased to 92.533% with TSA-CNN-AOA (Decision tree), to 95.007% with TSA-CNN-AOA (SVM) and to 95.098% with TSA-CNN-AOA (KNN).

Confusion matrices of all approaches are shown in Fig. 10. A confusion matrix is a widely used tabulation system that describes the prediction accuracy performance of a given model for each class label. In a confusion matrix, rows and columns correspond to the predicted class (Output Class) and true class (Target Class), respectively. It is clear from the matrix data that the classification accuracy rates of TSA-CNN-AOA (KNN) for negative, positive, and neutral Twitter data were 86.70%, 95.35%, and 97.64%, respectively. In addition, when all matrices are analyzed, it is evident that neutral Twitter data were classified with a higher accuracy rate, whereas negative data were classified with a lower accuracy rate. Receiver-operating characteristic (ROC) curves revealing the relationship between false positive rate (FPR) and true positive rate (TPR) are shown in Fig. 11.

In Table 2, the classification accuracy performance rate of the proposed approach was compared with other studies on SA of tweets about COVID-19 in the existing literature. It can be observed that the proposed approach displayed a higher classification performance compared to other proposed approaches in the current literature.

5 Conclusion

CNN has been a popular method for TSA in recent years. The present study, too, created a database consisting of tweets about COVID-19 for TSA to propose a new CNN-based hybrid approach. To this aim, tweets about COVID-19 were extracted from Twitter to create a large database and propose Twitter sentiment analysis using convolutional neural network optimized via arithmetic optimization algorithm (TSA-CNN-AOA). The proposed approach attempted to classify individuals' tweets about COVID-19 into three main categories: positive, negative, and neutral. Thus, it has become possible to reach significant conclusions about people's attitude toward the COVID-19 pandemic, which can help lessen and eliminate the impact of the disease on them. The experimental studies were performed to test the classification accuracy performances of TSA-CNN-AOA (Decision tree), TSA-CNN-AOA (SVM), and TSA-CNN-AOA (KNN) on the dataset, which yielded an accuracy rate of 92.533%, 95.007%, and 95.098%, respectively. Additionally, CNN, SVM, Naive Bayes, logistic regression, decision tree, and KNN approaches were also used for the testing process, and the highest classification accuracy rate was achieved by TSA-CNN-AOA (KNN). Finally, the classification performance of the proposed approach was compared with other proposed SA approaches in the current literature, indicating that the proposed approach displayed the highest performance. In conclusion, it can be stated that the present study proposes a remarkably successful approach for TSA.

Declarations

whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

1. WCOVID-19 Weekly Epidemiological Update Data as received by WHO from national authorities, as of 25 January 2022
2. Nasir A, Shah MA, Ashraf U, Khan A, Jeon G (2021) An intelligent framework to predict socio economic impacts of COVID-19 and public sentiments. Comput Electr Eng 96:107526
3. Nayak SR, Nayak DR, Sinha U, Arora V, Pachori RB (2021) Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. Biomed Signal Process Control 64:102365
4. Yadav M, Perumal M, Srinivas M (2020) Analysis on novel coronavirus (COVID-19) using machine learning methods. Chaos Solitons Fractals 139:110050
5. De Rosis S, Lopreite M, Puliga M, Vainieri M (2021) The early weeks of the Italian Covid-19 outbreak: sentiment insights from a Twitter analysis. Health Policy (Amsterdam, Netherlands)
6. Terpstra T, Stronkman RJP, de Vries A, Paradies GL (2012) Towards a realtime Twitter analysis during crises for operational crisis management. In: 9th international conference on information systems for crisis response and management, ISCRAM 2012, 22–25 April 2012, Vancouver, BC, USA. Simon Fraser University
7. Power R, Robinson B, Colton J, Cameron M (2014) Emergency situation awareness: twitter case studies. In: international conference on information systems for crisis response and management in mediterranean countries. Springer, Cham. (pp 218–231)
8. Zhou Z, Zhang X, Sanderson M (2014) Sentiment analysis on twitter through topic-based lexicon expansion. In: Australasian database conference. Springer, Cham. (pp 98–109)
9. Brynielsson J, Johansson F, Jonsson C, Westling A (2014) Emotion classification of social media posts for estimating people's reactions to communicated alert messages during crises. Secur Inform 3(1):1–11
10. Fiok K, Karwowski W, Gutierrez E, Wilamowski M (2021) Analysis of sentiment in tweets addressed to a single domain-specific Twitter account: comparison of model performance and explainability of predictions. Expert Syst Appl 186:115771
11. Go A, Huang L, Bhayani R (2009) Twitter sentiment analysis. Entropy 17:252
12. 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), pp 1320–1326
13. Kouloumpis E, Wilson T, Moore J (2011) Twitter sentiment analysis: the good the bad and the omg!. In: Proceedings of the international AAAI conference on web and social media (Vol. 5, No. 1, pp 538–541)
14. Xia R, Zong C, Li S (2011) Ensemble of feature sets and classification algorithms for sentiment classification. Inf Sci 181(6):1138–1152
15. Pagolu VS, Reddy KN, Panda G, Majhi B (2016) Sentiment analysis of Twitter data for predicting stock market movements. In: 2016 international conference on signal processing, communication, power and embedded system (SCOPES). Paralakhemundi, India. (pp 1345–1350)
16. Pota M, Esposito M, De Pietro G, Fujita H (2020) Best practices of convolutional neural networks for question classification. Appl Sci 10:4710
17. Tran K, Bisazza A, Monz C (2016) Recurrent memory networks for language modeling. arXiv preprint arXiv:1601.01272
18. Rao G, Huang W, Feng Z, Cong Q (2018) Lstm with sentence representations for document-level sentiment classification. Neurocomputing 38:49–57
19. Farhadloo M, Rolland E (2016) Fundamentals of sentiment analysis and its applications. In Sentiment analysis and ontology engineering. Springer, Cham. (pp 1–24)
20. Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau RJ (2011) Sentiment analysis of twitter data. In: Proceedings of the workshop on language in social media (LSM 2011) (pp 30–38)
21. Pang B, Lee L (2004) A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. arXiv preprint arXiv:cs/0409058
22. 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)
23. Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of human language technology conference and conference on empirical methods in natural language processing (pp 347–354)
24. Abdi A, Shamsuddin SM, Hasan S, Piran J (2019) Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. Inf Process Manag 56(4):1245–1259
25. Zhou S, Chen Q, Wang X (2014) Fuzzy deep belief networks for semi-supervised sentiment classification. Neurocomputing 131:312–322
26. Yadav N, Chatterjee N (2016) Text summarization using sentiment analysis for DUC data. In: 2016 international conference on information technology (ICIT), IEEE. (pp 229–234)
27. 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), IEEE. (pp 684–687)
28. Kalchbrenner N, Grefenstette E, Blunsom P (2014) A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188
29. Chen T, Xu R, He Y, Xia Y, Wang X (2016) Learning user and product distributed representations using a sequence model for sentiment analysis. IEEE Comput Intell Mag 11(3):34–44
30. Liu Y, Liu B, Shan L, Wang X (2018) Modelling context with neural networks for recommending idioms in essay writing. Neurocomputing 275:2287–2293
31. Tai KS, Socher R, Manning CD (2015) Improved semantic representations from tree-structured long short-term memory networks. arXiv preprint arXiv:1503.00075
32. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE Trans Signal Process 45(11):2673–2681
33. Kumar P, Vardhan M (2022) PWEBSA: twitter sentiment analysis by combining Plutchik wheel of emotion and word embedding. Int J Inf Technol 14(1):69–77
34. Villavicencio C, Macrohon JJ, Inbaraj XA, Jeng JH, Hsieh JG (2021) Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes. Information 12(5):204
35. Shamrat FMJM, Chakraborty S, Imran MM, Muna JN, Billah MM, Das P, Rahman OM (2021) Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. Indones J Electr Eng Comput Sci 23(1):463–470
36. Sontayasara T, Jariyapongpaiboon S, Promjun A, Seelpipat N, Saengtabtim K, Tang J, Leelawat N (2021) Twitter sentiment

analysis of Bangkok tourism during COVID-19 pandemic using support vector machine algorithm. J Disaster Res 16(1):24–30

37. Ankita A, Rani S, Bashir AK, Alhudhaif A, Koundal D, Gündüz ES (2022) An efficient CNN-LSTM model for sentiment detection in# BlackLivesMatter. Expert Systems with Applications, 116256

38. Usama M, Ahmad B, Song E, Hossain MS, Alrashoud M, Muhammad G (2020) Attention-based sentiment analysis using convolutional and recurrent neural network. Futur Gener Comput Syst 113:571–578

39. Behera RK, Jena M, Rath SK, Misra S (2021) Co-LSTM: convolutional LSTM model for sentiment analysis in social big data. Inf Process Manage 58(1):102435

40. Khasanah IN (2021) Sentiment classification using fasttext embedding and deep learning model. Procedia Comput Sci 189:343–350

41. Jain PK, Saravanan V, Pamula R (2021) A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. Trans Asian Low-Resour Lang Inf Process 20(5):1–15

42. Onan A (2021) Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. Concurr Comput Pract Exp 33(23):e5909

43. Jain D, Kumar A, Garg G (2020) Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN. Appl Soft Comput 91:106198

44. Nezhad ZB, Deihimi MA (2022) Twitter sentiment analysis from Iran about COVID 19 vaccine. Diabetes Metab Syndr 16(1):102367

45. Behl S, Rao A, Aggarwal S, Chadha S, Pannu HS (2021) Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. Int J Disaster Risk Reduct 55:102101

46. Basiri ME, Nemati S, Abdar M, Asadi S, Acharrya UR (2021) A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. Knowl-Based Syst 228:107242

47. Sitaula C, Basnet A, Mainali A, Shahi TB (2021) Deep learning-based methods for sentiment analysis on Nepali COVID-19-related tweets. Comput Intell Neurosci, 2021

48. AlBadani B, Shi R, Dong J (2022) A novel machine learning approach for sentiment analysis on Twitter incorporating the universal language model fine-tuning and SVM. Appl Syst Innov 5(1):13

49. Vernikou S, Lyras A, Kanavos A (2022) Multiclass sentiment analysis on COVID-19-related tweets using deep learning models. Neural Comput Appl, 1–13

50. Joulin A, Grave E, Bojanowski P, Mikolov T (2016) Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759

51. Ombabi AH, Ouarda W, Alimi AM (2020) Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks. Soc Netw Anal Min 10(1):1–13

52. Loria S (2018) Textblob Documentation. Release 0.15, 2, 269

53. Sohangir S, Petty N, Wang D (2018) Financial sentiment lexicon analysis. In: 2018 IEEE 12th international conference on semantic computing (ICSC), IEEE. (pp 286–289)

54. Ankit M, Saleena N (2018) An ensemble classification system for twitter sentiment analysis. Procedia Comput Sci 132(2):937–946

55. Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609

56. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830

57. Corinna C, Vladimir V (1995) Support-vector networks. Mach Learn 20(3):273–297

58. Bayes T (1968) Naive bayes classifier. Article Sources and Contributors, 1–9

59. Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theory 13(1):21–27

60. Surya Prasath VB, Arafat Abu Alfeilat H, Hassanat ABA, Lasassmeh O, Tarawneh AS, Bashir Alhasanat M, Eyal Salman HS Effects of distance measure choice on KNN classifier performance—A review. arXiv 2017, arXiv:1708.04321

61. Aksu G, Dogan N (2019) Comparison of decision trees used in data mining= Veri madenciliginde kullanilan karar agaçlarinin karsilastirilmasi. Pegem J Educ Instr 9(4):1183–1208

62. Investor fact sheet. Twitter. 2021. [29–04–2021] https://s22.q4cdn.com/826641620/files/doc_financials/2021/q1/Q1'21-Shareholder-Letter.pdf

63. We used MAXQDA 2020 (VERBI Software, 2019) for data analysis

64. Naseem U, Razzak I, Khushi M, Eklund PW, Kim J (2021) COVIDSenti: a large-scale benchmark Twitter data set for COVID-19 sentiment analysis. IEEE Trans Comput Soc Syst 8(4):1003–1015

65. Nair AJ, Veena G, Vinayak A (2021) Comparative study of twitter sentiment on covid-19 tweets. In: 2021 5th international conference on computing methodologies and communication (ICCMC), IEEE. (pp. 1773–1778)

66. Al-Sarem M, Alsaeedi A, Saeed F, Boulila W, AmeerBakhsh O (2021) A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs. Appl Sci 11(17):7940

67. Khakharia A, Shah V, Gupta P (2021) Sentiment analysis of COVID-19 vaccine tweets using machine learning. Available at SSRN 3869531

68. Jalil Z, Abbasi A, Javed AR, Khan MB, Hasanat MHA, Malik KM, Saudagar AKJ (2021) COVID-19 related sentiment analysis using state-of-the-art machine learning and deep learning techniques. Frontiers in Public Health, 9

69. Rustam F, Khalid M, Aslam W, Rupapara V, Mehmood A, Choi GS (2021) A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. PLoS One 16(2):e0245909