



# Machine learning-based IoT system for COVID-19 epidemics

Micheal Olaolu Arowolo<sup>1</sup> · Roseline Oluwaseun Ogundokun<sup>2</sup> · Sanjay Misra<sup>3</sup> · Blessing Dorothy Agboola<sup>4</sup> · Brij Gupta<sup>5,6,7</sup>

Received: 11 April 2021 / Accepted: 13 December 2021 / Published online: 1 March 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

## Abstract

The planet earth has been facing COVID-19 epidemic as a challenge in recent time. It is predictable that the world will be fighting the pandemic by taking precautions steps before an operative vaccine is found. The IoT produces huge data volumes, whether private or public, through the invention of IoT devices in the form of smart devices with an improved rate of IoT data generation. A lot of devices interact with each other in the IoT ecosystem through the cloud or servers. Various techniques have been presented in recent time, using data mining approach have proven help detect possible cases of coronaviruses. Therefore, this study uses machine learning technique (ABC and SVM) to predict COVID-19 for IoT data system. The system used two machine learning techniques which are Artificial Bee Colony algorithm with Support Vector Machine classifier on a San Francisco COVID-19 dataset. The system was evaluated using confusion matrix and had a 95% accuracy, 95% sensitivity, 95% specificity, 97% precision, 96% F1 score, 89% Matthews correlation coefficient for ABC-L-SVM and 97% accuracy, 95% sensitivity, 100% specificity, 100% precision, 97% F1 score, 93.1% Matthews correlation coefficient for ABC-Q-SVM. In conclusion, the system shows that the process of dimensionality reduction utilizing ABC feature extraction techniques can boost the classification production for SVM. It was observed that fetching relevant information from IoT systems before classification is relatively beneficial.

**Keywords** COVID-19 · ABC · SVM · Machine learning · IoT

**Mathematics Subject Classification** 68

## 1 Introduction

The novel coronavirus is a contributing agent (COVID-19), believed to originate from a China province, called Wuhan in late 2019 [1–6]. There is no clinically proven vaccine or drugs in the new coronavirus (COVID-19) contagion. It is evident that non-clinical

methods such as machine learning methods, internet of things, improved intelligence and artificial intelligence methods, the universe requires a quicker and faster resolution to include and address futuristic spread of COVID-19 worldwide in order to alleviate the enormous load on the healthcare organization to give the top conceivable means for the analysis of patients [7]. It is apparent that non-clinical assistant methods such as machine learning methods and improved intelligence such as the internet of things, the world desires a faster means to regulate and resolute more COVID-19 spread worldwide in order to alleviate the immense strain on the medical industries despite making available the greatest likely mediums for diagnosing individuals [8]. Machine learning is an innovative technique of artificial intelligence for discovering innovative, beneficial, and true unseen patterns or dataset. In multiple or single datasets, the method exposes relations and information amongst datasets [9]. A valuable source for mining and analysis of valuable, effective, and innovative information removal for improved decision-making to include the epidemic of the COVID-19 is a huge dataset created worldwide related to the 2019-nCoV pandemic every day. Machine learning is commonly used approach in diverse healthcare field applications, for example forecasting patient performance, demonstrating clinical results, health sector rating, and evaluating the efficacy of treatment as well as contamination prevention, steadiness, and retrieval [10]. There has been tremendous growth in healthcare intensive care systems in hospitals as well as numerous health hubs. Transferable healthcare nursing schemes exist with innovative knowledge. They have become boundless worry to various nations through the world. The initiation of Internet of Things (IoT) technology from direct consultation to telemedicine promotes the development of healthcare [7]. In the last decade, IoT has made all artifacts internally connected, and it has been seen as the next technological breakthrough. Some of the IoT apps include smart health monitoring mechanisms, smart home, smart parking, smart climate, smart city, industrial spots, and farmed grounds. The most comprehensive application of IoT is in health management, which delivers monitoring services for health and environmental conditions. Using sensors and networks, IoT is nonentity but linking supercomputers to the cyberspace [10]. The IoT data system of COVID-19 will perform a good function in the discovery of infected and uninfected individuals. For all countries and countries facing many difficulties in identifying COVID-19 patients, COVID-19 has become a pandemic problem because its symptoms are not easy to detect and are spread by contacting each other [11]. This current research comprises the following important research issues:

- (i) What are some new Artificial Intelligence, machine learning, and Internet of Things applications which can aid in identifying risks against the COVID-19 infection?
- (ii) What information from social science, Artificial Intelligence, and machine learning applications played an important role throughout this conflict?
- (iii) What significance do development innovations like the Internet of Things play in COVID-19 treatment and elimination?

In this study, a machine learning process is carried out to predict COVID-19 cases using Artificial Bee Colony algorithm with Support Vector Machine classifier using a San Francisco COVID-19 dataset. Thus, artificial intelligence, the Internet of Things,

and machine learning open up new mechanisms for successful healthcare throughout that epidemic, which can be valuable for pharmaceutical creation, efficient diagnosis tactics, and disease spreading predictions. The remaining part of this article is arranged as Sect. 2 discussing the literature review, Sect. 3 discussing the material and methodology, Sect. 4 discussing the proposed system results and discussion with the system evaluation and comparison with previous researches. Section 5 concluded the study with a conclusion and recommendations.

## 2 Literature reviews

Machine learning and IoT has been employed in a different field such as information and network security [12–15], agricultural sector [16–20], Medical or healthcare industry [21–24] and so on. Many researchers are encouraged to establish an automatic diagnostic system because improved premature cancer detection can increase patients' survival rates. In this field of research, the number of publications is increasing exponentially and work is still underway to show significant progress towards the performance of an efficient and robust segmentation tool. This section presents the related findings of research in previous years. An IoT smart health monitoring approach was proposed by researcher [10], for tracking the elementary health signs of patients and the situation in the area where patients are situated in actual period. Five sensors were executed to collected the heartbeat, body and area temperature, CO with CO2 sensors information from the hospital area. For each case, the fault fraction of the system produced is in a positive boundary ( $< 5\%$ ). The patients' state is transmitted to the medical staff through a portal, processed and interprets the patients' present condition. For healthcare monitoring, which is demonstrated by the efficacy of the device, the built prototype is well suited. Machine learning models was developed with an epidemiology dataset of patients to predict COVID [25]. The decision tree, naive Bayes, support vector machine, logistic regression, K-NN and random forest algorithms were utilized. For patients to improve medically, the model predicted days, patients age group with high danger or not recovering from the pandemic, those likely to recover and those probable to improve fast from the disease. The outcomes showed that the model built with the decision tree algorithm is well competent in predicting the probability of regaining with an overall accuracy of 99.85%. An overview of healthcare monitoring system based on IoT has been carried out by researcher [26]. The findings of the study relate to the use of IoT in healthcare facilities, helping to achieve correct medical diagnoses, focusing on standard service given to patients. Also decreases intermittent hospital patient evaluations by depending on remote diagnosis IoT applications. An application to health institutions would also lead to yielding accurate information for the ailments suffered by patients and thereby engaging them in the preparation of clinical studies to achieve more detailed outcomes. This study presented the Internet-based healthcare nursing classification. Neural Network, Support Vector Machine, Naïve Bayes, Decision tree, KNN, OneR, with ZeroR algorithms was proposed as machine learning approaches by researcher [27], they tested these algorithms, after selecting the appropriate symptoms, on COVID-19 dataset. Five algorithms realized an accuracy of about 90%. The study assume that the data will permit

these 5 algorithms to classify latent COVID-19 cases efficiently and accurately. The system will record the answer to treatment for respective patients that contracted the virus. Covid-19 is a pandemic spreading through the biosphere. Scientists and engineers are working to develop a serum for further investigative services and strengthen monitoring approaches. Monitoring the individual's health, mobile and web-based applications, based on questionnaires, have already been developed. Preventing the blowout of Covid-19, requires approaches such as the Internet of Things (IoT). It is an interrelation of the Internet and corporal devices. This paper reviewed the available Covid-19 surveys, monitoring techniques, and proposed an IoT-based planning used for minimizing blowout of Covid-19 [28]. The most recent information on artificial intelligence for COVID-19 was gathered and then investigated to recognize its potential disease application. Seven substantial AI applications have been reported for the COVID-19 contagion. This technology delivers a significant responsibility in defining the case cluster and predicting the location the disease will disrupt in the future through the compilation and review of all preceding data. Medical establishments are in desperate requirement of technology for administrative to deal with COVID-19 and to assist them in having appropriate real-time submissions to prevent its transmission. AI works to imitate human intellect in a professional way. In recognizing and recommending the development of a COVID-19 vaccine, it might perform a substantial responsibility as well [29]. Exploitation of SVM Radial, SVM Polynomial, and Random Forest techniques, likened to every further method utilized in the investigation, a machine learning method created an ensemble voting classifier that provides greater precision, accuracy, specificity, recall, and F1 score. The proposed ensemble model has predicted a sum of 1326 probable human aimed proteins of SARS-CoV-2 and authenticated them utilizing gene ontology and KEGG passage enhancement examination. Numerous reprovable medications are already known that target the predicted interactions. This research will facilitate the detection of future targets for the more productive discovery of anti-COVID drugs [30]. For reception by SARS-CoV-2 PCR research, which had a number of auxiliary test center constituents as well ( $n = 1455$ ) in 2020, all emergency rooms or inpatient cases were recommended. For the final diagnostic classification, they checked 7 machine learning methods and utilized a mixture of those methods. The joint methods had an extent under the receiver operator curve of 0.91 (95 percent confidence interval 0.87–0.96) in the test set ( $n = 392$ ). The model obtained a 0.93 (95 percent CI 0.85–0.98) sensitivity and a 0.64 (95 percent CI 0.58–0.69) specificity. Likened to the SARS-CoV-2 PCR, the authors discovered that their projected machine learning (ML) algorithm had an outstanding analytical system of measurement. This collection of ML algorithms to establish COVID-19 can be utilized in infirmary settings where PCR examination is limited or inaccessible as a screening tool [31]. A DCNN model was suggested to alter the arithmetical characteristics from the innovative dataset to the image domain and then fill them into a DCNN method. 3 convolutional layers, 3 ReLU layers, 3 pooling layers, and 2 completely related layers are used in the proposed DCNN model. The experimental results indicate that as compared to traditional machine learning, for instance, support vector machine, decision tree, and ensemble, the suggested DCNN method for management classification (32 classes) attained 98.05 per cent in research correctness. The conventional machine learning (ensemble) algorithm attained 98.5 percent in study precision in the calculation of

treatment concentration level, whereas the projected DCNN method attained 98.2 percent. The performance metrics affirm the accuracy of the outcomes obtained from the research carried out in the classification of therapy and the prediction level of treatment focus [32]. Utilizing traditional and ensemble machine learning algorithms, it has been suggested to divide documented medical reports into 4 categories. Function engineering has been achieved utilizing approaches such as term occurrence/converse text frequency (TF/IDF), bag of words (BOW) and study extent. With these characteristics, conventional and ensemble machine learning classifiers were fitted. Logistic regression and Multinomial Naïve Bayes displayed greater outcomes than further ML procedures by having 96.2 percent test correctness. It was suggested that recurrent neural networks can be utilized in the future with greater accuracy [33]. A relative analysis of machine learning algorithm to forecast the outbreak of COVID-19 has been projected. In the 2 algorithms, auspicious outcomes were seen (that is multifaceted perceptron, MLP, and adaptive network-based fuzzy inference system, ANFIS). Based on the results released, and due to the extremely dynamic nature of the COVID-19 outbreak and the variation in its behavior from nation to nation, this study indicates machine learning as an effective tool to model the outbreak. This paper provides an initial benchmark to highlight the promise of machine learning for future research. Furthermore the paper indicates that by combining machine learning and SEIR models, real innovation in outburst estimation can be comprehended [34]. Artificial intelligence investigators are concentrating their proficiency on developing mathematical methods using nationally shared data to analyze this epidemic situation. This article proposes the use of machine learning and deep learning models to devote to the welfare of residing in a the world to understand its daily exponential behavior and to predict the future accessibility of COVID-2019 across nations by using the Johns Hopkins dashboard's real-time data [35]. An amalgam machine learning strategy was proposed to envisage COVID-19 and illustrate its possibility of utilizing Hungarian data. The adaptive network-based fuzzy inference system (ANFIS) and multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) hybrid machine learning methods are projected to envisage the time series and mortality rate of disease-ridden persons. The system forecasted that the outburst and overall ethics will fall considerably by late May. The authentication is done with promising results for 9 days, which confirms the accuracy of the model. The model is expected to retain its precision in as much as there is no momentous disruption occurring. In order to show the probability of machine learning for impending investigation, the article provides an initial benchmarking [36]. In author [32], they concentrate on current advances in the expansion of artificial intelligence-based COVID-19 medications and inoculations and the ability for smart training to determine COVID-19 infected individual. To promote deep learning applications for SARS-COV-2, several molecular targets of COVID-19, whose inhibition can improve patient survival, are illuminated. In accumulation, existing Corona-DB-AI, a dataset of in silico or in vitro discovered molecules, peptides, and epitopes that can theoretically be utilized to derive COVID-19 care for training models. The knowledge and datasets presented in this analysis can be used to train deep learning-based models and accelerate the discovery of efficient viral therapies [37]. Machine learning algorithms are able to evaluate a large number of parameters to determine the predictors of disease outcomes within a limited period of time. It would

be useful in decision-making for clinicians to incorporate such an algorithm to predict high-risk individuals during the early stages of infection in order to avoid permanent harm. Here, we propose suggestions for the development of prognostic machine learning models using electronic health records so that COVID-19 can develop a real-time risk score [38]. Many novel associations have been identified between clinical factors, including links between being male and possessing higher numbers of serum lymphocytes and neutrophils. We observed that COVID-19 patients could be grouped into subtypes based on serum immune cell numbers, sex, and recorded symptoms. Finally, we have trained an XGBoost model to achieve a sensitivity of 92.5 percent and a precision of 97.9 percent in separating COVID-19 patients from influenza patients. We demonstrated that computer techniques trained on massive clinical databases could create ever more reliable COVID-19 diagnostic models to mitigate the effect of lack of research. For previously unknown COVID-19, we have given clinical component associations and clinical subgroups [39]. COVID-19 causes immunity, and if so, we do not know for how long whether "herd immunity" can potentially mitigate the risk or if an appropriate vaccine can be made, and this knowledge may be a long time coming. The COVID-19 pandemic, meanwhile, poses enormous obstacles for medical research and clinical trials in particular. In this article, some of these problems are highlighted and propose ways in which machine learning can help respond to those challenges. Three challenge areas were identified: existing clinical trials of non-COVID-19 drugs; clinical trials of COVID-19 presupposition drugs; and clinical trials of experimental COVID-19 drugs [40]. In summary, Current findings have suggested that proposed models for COVID-19 identification based on algorithms can improve the accuracy of Covid-19 diagnosis. Machine Learning and Artificial Intelligence may be used to foresee and predict the new epidemic, according to selective information. With the novel methodology, a novel prediction and prediction model for Covid-19 patients has been developed.

### 3 Materials and methods

The dataset was obtained from the host transcriptional response to SARS-CoV-2 in 234 patients with COVID-19 ( $n = 93$ ), other viral ( $n = 100$ ) or non-viral ( $n = 41$ ) acute respiratory diseases (ARIs) through metagenomic sequencing of upper airway samples. COVID-19 was characterized by a decreased innate immune response compared to other viral ARIs, with decreased expression of genes involved in toll-like signaling of receptors and interleukins, chemokine binding, neutrophil degranulation, and lymphoid cell interactions. Compared to other viral ARIs, patients with COVID-19 also exhibited significantly reduced proportions of neutrophils, macrophages, and increased proportions of goblet, dendritic, and B cells. We constructed 27-, 10- and 3-gene classifiers using machine learning to differentiate COVID-19 from other acute respiratory diseases [41].

MATLAB was used to analyze the data collected from an experimental method, and ABC was used to select features. The selected features were used to use the SVM algorithm approach to carry out classification.

### 3.1 Artificial bee colony

This study presents an enhanced ABC algorithm for fetching relevant genes from COVID-19 transcriptional data. The Artificial Bee Colony algorithm is one of the most recent nature-inspired optimization algorithms based on the intelligent foraging behavior of the honey bee swarm. Karaboga suggested and further developed the ABC algorithm. The ABC algorithm showed impressive results for a wide variety of concerns [42]. Three categories are classified into foraging honey bees, namely, working bees, onlooker bees and scout bees. One unique operation for producing new candidate solutions is symbolized by each group of honey bees. Employed bees exploit food supplies. From different food sources, they add nectar to their hive. Onlooker bees wait in the hive for information on food sources to be shared by working bees and search for a food supply based on that information. Becoming scouts and abandoning their solutions are working bees whose food resources have been exhausted. Then the scout bees look randomly for new food sources near the hive without using any information. After the scout finds a new food source, it again becomes an employed bee. Each scout is an adventurer who has no direction when looking for a new food, i.e. a scout might find some sort of source of food. Therefore a scout may unwittingly discover a more affluent and entirely unknown food source [43].

#### The procedure of the ABC algorithm

Starting phase

- 1- Set the samples and allocate the samples to active bees
- 2- While (sequence = MAX\_Sequence) do
  - Active bee Stage
  - 3- for  $i = 1$  to SN do
  - 4- Generate new result  $v_i$  for active bees and analyze fitness rate
  - 5- Introduce greedy assortment device among  $v_i$  and  $x_i$ , pick best choice
  - 6- If result  $x_i$  does not update, the non-updated number  $t_i = t_i + 1$ ; else  $t_i = 0$
  - 7- end for
  - Observer bee stage
  - 8- Analyze the selection probability  $p_i$
  - 9-  $t = 0, I = 1$
  - 10- while ( $t < SN$ ) do
  - 11- if random  $< P_i$
  - 12-  $t = t + 1$
  - 13- Generate a result for observer bee of the result and analyze its fitness rate
  - 14- Use greedy selection approach between  $v_i$  and  $x_i$ , pick better one
  - 15- If result  $x_i$  does not update, the non-updated number  $t_i = t_i + 1$ ; else  $t_i = 0$
  - 16- end if
  - 17-  $i = i + 1$
  - 18- If  $i = SN + 1, i = 1,$
  - 19- end while ( $t = SN$ )
- Detect bee Stage

- 20-  $t_i > \text{limit}$  then
- 21- Change  $x_i$  with new random result
- 22- end if
- 23- Learn present optimal result
- 24- sequence = sequence + 1
- 25- while (sequence = MAX\_Sequences)

New solutions are developed using a neighborhood operator in the employed bee and onlooker bee process. In order to improve the exploitation potential of ABC, a local search procedure is applied to the solution obtained by the neighborhood operator with a certain probability. In addition to the approach implemented, we have further updated it by incorporating two new components into it in order to resolve the shortcomings of the ABC algorithm. Second, the pheromone idea, which is one of the key components of the Ant Colony Optimization (ACO) algorithm, has been implemented [43].

### 3.2 Support vector machine

The Support vector machine (SVM) is one of the supervised learning algorithms used for classification and regression. The classification activity in SVM involves testing and training data containing those data instances. In the training dataset, each instance includes one or more target values; thus, the primary aim of SVM is to produce a model that predicts target value or values. SVM proposed the introduction of alternate regression loss functions, which can be linear or nonlinear [44].

SVM is one of the most popular and impressive machine learning approaches for recognition and regression. SVM where 1 is the mark from 0 to 1,  $w$  is a working algorithm. The output is  $a-q$ , the coefficients of linear group are  $w$  and  $q$ , and the input vector is a [45].

$$\begin{aligned}
 SVM_{hk} &= \max(0, 1 - |k(w \cdot ak - q)|) \\
 SVM_{loss} &= \sum t = \max(0, ht) \\
 v(x) &= \text{sign} \left[ \sum k = \ln \alpha k y k \phi(x, xk) + b \right]
 \end{aligned}$$

where  $\alpha k$  = positive real constant,  $b$  = real constant

### 3.3 Performance evaluation

To determine their accuracy, computer models are tested using assessment techniques. The techniques use data mining algorithms or machine learning algorithms to assess the consistency and efficacy of the model. These key performance assessment techniques for the data mining model include precision, sensitivity, and accuracy. In this analysis, however, to evaluate the established models, the only accuracy is considered [7]. Some validation metrics include assessing the efficiency of the machine learning model. In classification models, the uncertainty matrix is often used to evaluate four characteristics: True Positive (TP), False Positive (FP), True Negative (TN) and False

Negative (FN). The correctly and incorrectly classified illustrations from the sample of the dataset provided to test the model are discovered [46].

## 4 Applications

The need to explore similar genes is beneficial in developing different applications such as updated therapy, cancer detection, gene and drug development, tumor classification, diseases which include typhoid, malaria, and others. Machine learning technique has outstanding algorithms as approaches that are applied to innumerable areas when discovering the designs and the difference between data. MATLAB (Matrix Laboratory) is used to execute the research because of it has simple and advantageous programming setting for technologists, designers, physicists, scholars, amid others. The MATLAB arithmetical multi-worldview is a processing setting and restricted programming language developed by MathWorks. It allows machine monitoring, plotting of functions and information, implementation of algorithms, generation of user interfaces, inscribed in diverse languages for instance, C, C++, C#, Java, Fortran and Python [47–49]. The key aim of this investigation is to use the COVID-19 database to predict the use of the MATLAB system. The iCore2 cpu, 4 GB RAM size, 64-bit scheme and MATLAB 2015a are used by the computer configuration as the execution tools for testing this learning.

## 5 Results and discussion

This study reveals that with 235 properties, the novelty of COVID-19 includes 15979 instances of associated genes. The ABC approached was applied on the data to decrease the curse of dimensionality. ABC dimensionality reduction senses and obtains similar characteristics to determine optimum variation for a reduced number of sub-set characteristics. In this report, ABC is added to the COVID-19 results, which provides important gene information which is beneficial for more research. Classification algorithms connect SVM kernels using the MATLAB approach to the model's implementation. Using ABC as a technique to decrease the dimensionality of feature collection, unique subsets of feature genes were selected with substantial records. SVM classification kernels with 10-fold cross validation were used to determine the execution of the utility of the classification methods, utilizing 0.05 constraint holdout data for training and 5% for testing to determine the correctness of the classifiers. The classifier follows a learning valuation protocol, the planning and research processes are estimated as a 10-fold cross validation to remove the selection biases. This protocol is being executed utilizing MATLAB. Computational time and efficiency measurements (Accuracy, Precision, Sensitivity, Accuracy, F-score and Recall) are the basis of the recorded outcome of the valuation. Figure 1 shows the COVID-19 datasets loaded.

The loaded data is passed into the ABC algorithm and the output is shown in Fig. 2.

This study uses ABC in the loaded data to fetch the sub-set of relevant features. The features selected are passed into the classification of the SVM and the result is

shown in Figs. 5, 6. A solution to the performance metrics is given by the confusion matrix. Figures 3 and 4 shows the scattered plot and ROC curve for the experiment respectively (Figs. 5, 6).

A COVID-19 data [41] with 15979 gene function was used as a dimensionality reduction method to evaluate the efficiency of the machine learning scheme to obtain the related features.

To forecast their efficiency, these components are then categorized using SVM classification. The outcome illustrates the effectiveness of genes in machine learning technology. The performance outcomes are revealed and likened in Table 1 to support the strategy. The outcome demonstrates that KNN does better than Decision Tree in respect to a reduced amount of preparation time and output precision. In this study, COVID epidemic analysis was carried out using ABC with SVMs. This analysis examined and enhanced the classification of COVID-19 data, numerous studies were suggested by researchers in reviews utilizing the performance metrics revealed in Table 1, the outcomes showed that the process of dimensionality reduction utilizing ABC feature extraction techniques can boost the classification production for SVM. In this study, the result of the study compares with other novel studies, as shown in Fig. 6. It was observed that several studies need to introduce approach such as dimensionality reduction in fetching relevant information that can benefit the classification. It was observed that fetching relevant information before classification is relatively beneficial. Running time was also employed to evaluate the performance of the machine learning classification techniques employed in this study and it was discovered that the

235 Attributes loaded      15979 Instances loaded

GSE156063\_swab\_gene\_counts.

1	1534	621	1903	2796	416	15	^
1	153	123	221	344	112	20	
1	295	185	457	626	170	79	
1	75	108	292	367	49	31	
1	19	471	1843	284	38	539	
1	500	288	687	1528	311	20	
1	271	190	655	663	98	23	
1	2463	2540	3969	11399	1073	91	
1	143	167	637	345	136	87	
1	380	337	1209	816	209	12	
1	574	304	1017	1588	219	18	
1	217	108	553	420	146	21	
1	772	707	1513	2647	315	42	
1	310	154	695	1409	175	16	
1	145	255	549	321	64	6	
1	594	473	1310	1439	808	108	
1	158	207	417	331	128	59	
1	290	338	637	833	134	14	
1	06	120	200	154	21	22	^

SAVE

Fig. 1 Loaded COVID-19 data

52 Attributes loaded		3013 Instances loaded				
GSE1560634x4.xlsx						
1	1534	621	1903	2796	416	15 ^
1	153	123	221	344	112	20
1	295	185	457	626	170	79
1	75	108	292	367	49	31
1	19	471	1843	284	38	539
1	500	288	687	1528	311	20
1	271	190	655	663	98	23
1	2463	2540	3969	11399	1073	91
1	143	167	637	345	136	87
1	380	337	1209	816	209	12
1	574	304	1017	1588	219	18
1	217	108	553	420	146	21
1	772	707	1513	2647	315	42
1	310	154	695	1409	175	16
1	145	255	549	321	64	6
1	594	473	1310	1439	808	108
1	158	207	417	331	128	59
1	290	338	637	833	134	14
1	66	120	200	154	21	??

Fig. 2 Reduced data using ABC

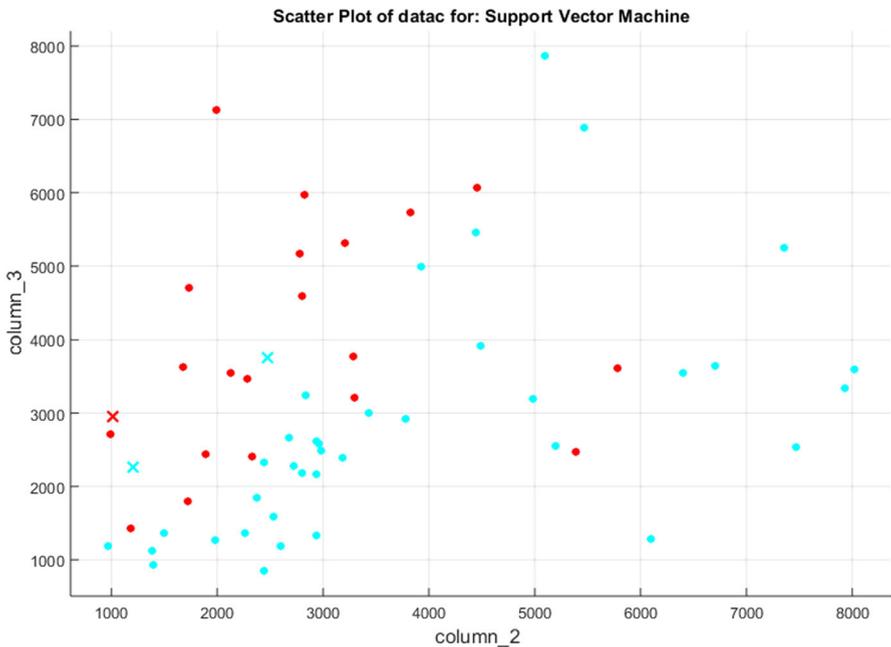


Fig. 3 ABC reduced COVID-19 data scattered plot

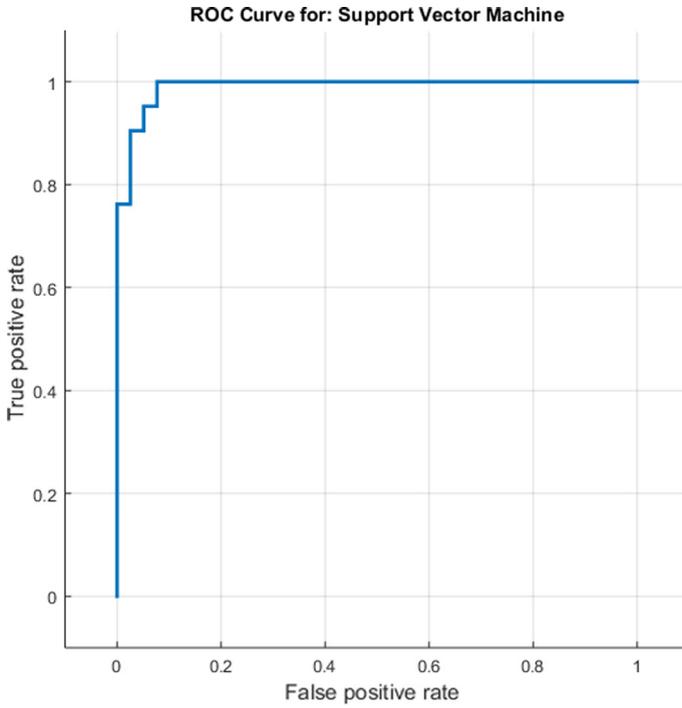


Fig. 4 The experiment ROC curve

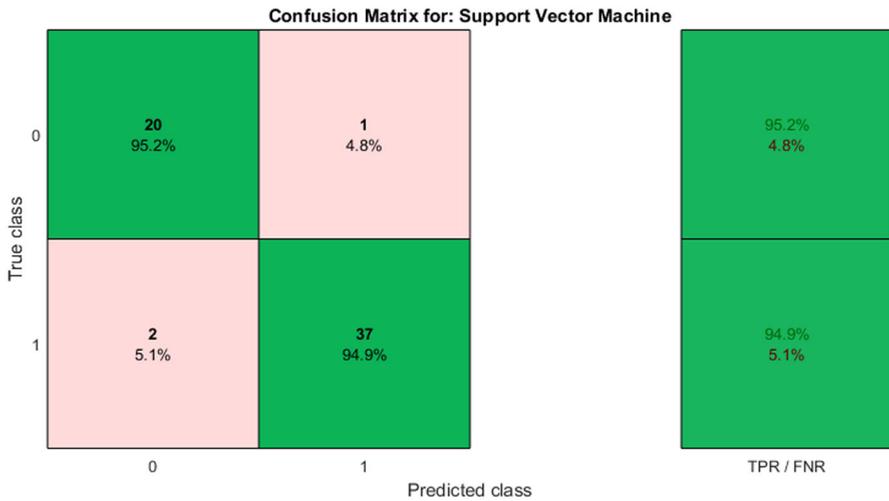
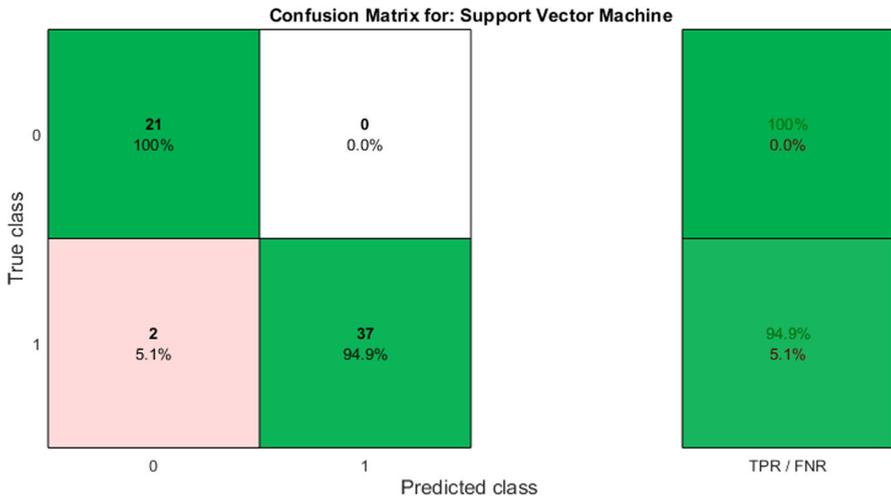


Fig. 5 Complete accuracy for and confusion matrix for the classification of COVID-19 data using L-SVM TP=37; TN=20; FP=1; FN=2



**Fig. 6** Overall accuracy for and confusion matrix for the classification of COVID-19 data using quadratic-SVM TP=37; TN=21; FP=0; FN=2

**Table 1** Performance metrics table for the confusion matrix

Performance metrics	ABC-L-SVM classification	ABC-Q-SVM classification
Accuracy (%)	95	97
Sensitivity (%)	95	95
Specificity (%)	95	100
Precision (%)	97	100
F1 score (%)	96	97
Matthews correlation coefficient (%)	89.0	93.1

**Table 2** Comparison with existing works

Authors	Approaches	Results (%)
Muhammad et al. [12]	DT	99.85
Li et al. [16]	XG-Boost	92.5
Goodman [17]	ML	93
Khalifa [19]	DCNN	98
Khanday [20]	Multinomial Naïve Bayes	96

system had a superlative running time of 5.34 mins for the training phase and 0.076 secs for the detection time (Table 2).

## 6 Conclusion and future work

This research strengthens and can be successful in the prognosis and diagnosis of COVID-19. The recommended solution utilized machine learning methods for instance model reduction of dimensionality and classification algorithms. The function collection ABC model for the SVM classifiers uses the dimensionality reduction model. This thesis carried out the performance analysis and assessment and shows the results obtained, Q-SVM outperforms the algorithm of L-SVM classification. This analysis examined and enhanced the classification of COVID-19 data, several studies were suggested by researchers in reviews using performance metrics, the findings showed that the model of dimensionality reduction would help boost classification production. IoT databases can help in the monitoring and evaluation of COVID-19 infections through gene features. If recent proposed studies will develop feature extraction models and algorithms, it will be important to investigate. This research confirms the need for Artificial Intelligence, Machine Learning, and Internet of Things (IoT) technologies for fending against the COVID-19 pandemic with futuristic work conducting more robust data as well as conducting an experiment on classifier such as KNN. Integration into an IoT scheme of the advantages of telemedicine, such as the possibility of it would be possible to access health records through software and help electronic communication networks between doctor and patient. Upgrade the system to one that is ideal for long-term use of the medical system, not just for our immediate crisis. Early detection of instances using the suggested approach could potentially lessen the burden of infectious diseases as well as mortality rates. This approach would also allow for the tracking of recovered instances as well as improved management of the importance. In the future, the authors can adopt the use of deep learning algorithms for the implementation and detection of the COVID-19 disease. The authors can also employ the use of visual saliency for detection in IoT systems. They can also employ the use of compound rank-k projection for the analysis of the study to acquire a robust feature in the IoT system to get a better result. Self-supervised learning can also be employed in the future.

## References

1. Lukman AF, Rauf RI, Abiodun O, Oludoun O, Ayinde K, Ogundokun RO (2020) COVID-19 prevalence estimation: four most affected African countries. *Infect Dis Model* 5:827–838
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, Xing X (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 382:1199–1207
3. Ogundokun RO, Lukman AF, Kibria GBM, Awotunde JB, Aladeitan BB (2020) Predictive modelling of COVID-19 confirmed cases in Nigeria. *Infect Dis Model* 5:543–548
4. Wolfel R, Corman V, Guggemos W, Seilmaier M, Zange S, Müller M et al (2020) Virological assessment of hospitalized cases of coronavirus disease. *Nature* 581:465–469
5. Daniyal M, Ogundokun RO, Abid K, Khan D, Ogundokun OE (2020) Predictive modeling of COVID-19 death cases in Pakistan. *Infect Dis Model*
6. Mahase E (2020) Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite a lower case fatality rate
7. Islam MM, Rahaman A, Islam MR (2020) Development of smart healthcare monitoring system in IoT environment. *SN Comput Sci* 1:1–11

8. Debnath S, Barnaby DP, Coppa K, Makhnevich A, Kim EJ, Chatterjee S et al (2020) Machine learning to assist clinical decision-making during the COVID-19 pandemic. *Bioelect Med* 6(1):1–8
9. Jibril ML, Sharif US (2020) Power of artificial intelligence to diagnose and prevent further covid-19 outbreak: a short communication. arXiv preprint [arXiv:2004.12463](https://arxiv.org/abs/2004.12463)
10. Rahaman A, Islam MM, Islam MR, Sadi MS, Nooruddin S (2019) Developing IoT based smart health monitoring systems: a review. *Rev. dIntelligence Artif.* 33(6):435–440
11. Alaba FA, Othman M, Hashem IAT, Alotaibi F (2017) Internet of things security: a survey. *J Netw Comput Appl* 88:10–28
12. Cvitić I, Peraković D, Periša M, Gupta B (2021) Ensemble a machine learning approach for the classification of IoT devices in a smart home. *Int J Mach Learn Cybern* 1–24
13. Cvitić I, Peraković D, Periša M, Husnjak S (2019) An overview of distributed denial of service traffic detection approaches. *Promet Traffic Transp* 31(4):453–464
14. Abdulsalam SO, Mohammed AA, Ajao JF, Babatunde RS, Ogundokun RO, Nnodim CT, Arowolo MO (2020) Performance Evaluation of ANOVA and RFE Algorithms for Classifying Microarray Dataset Using SVM. In: Themistocleous M, Papadaki M, Kamal MM (eds) *Information systems. lecture notes in business information processing*, vol 402. Springer, Cham. [https://doi.org/10.1007/978-3-030-63396-7\\_32](https://doi.org/10.1007/978-3-030-63396-7_32)
15. Cvitić I, Peraković D, Periša M, Botica M (2021) A novel approach for the detection of IoT-generated DDoS traffic. *Wireless Netw* 27(3):1573–1586
16. Ogundokun RO, Awotunde JB, Misra S, Abikoye OC, Folarin O (2021) Application of machine learning for ransomware detection in IoT devices. *Artificial intelligence for cyber security: methods, issues and possible horizons or opportunities*. Springer, Cham, pp 393–420
17. Nnodim CT, Arowolo MO, Agboola BD, Ogundokun RO, Abiodun MK (2021) Future trends in mechatronics. *Int J Robot Autom* 10(1):24–31
18. Devi YSS, Prasad TKD, Saladi K, Nandan D (2020) *Analysis of precision agriculture technique by using machine learning and IoT. Soft computing: theories and applications*. Springer, Singapore, pp 859–867
19. Nnodim CT, Fath El-Bab AMR, Ikua BW, Sila DN (2019) Estimation of the modulus of elasticity of mango for fruit sorting. *Int J Mech Mechat Eng* 19(2):1–10
20. Nnodim CT, Fath El-Bab AMR, Ikua BW, Sila DN (2021) Design, simulation, and experimental testing of a tactile sensor for fruit ripeness detection. In: Ao SI, Kim HK, Amouzegar MA (eds) *Transactions on engineering technologies*. Springer, Singapore, [https://doi.org/10.1007/978-981-15-9209-6\\_5](https://doi.org/10.1007/978-981-15-9209-6_5)
21. Adenugba F, Misra S, Maskeliūnas R, Damaševičius R, Kazanavičius E (2019) Smart irrigation system for environmental sustainability in Africa: an internet of everything (IoE) approach. *Math Biosci Eng* 16(5):5490–5503
22. Pandey PS (2017) Machine learning and IoT for prediction and detection of stress. In: *Proceedings of the 2017 17th international conference on computational science and its applications (ICCSA)* (pp 1–5). IEEE
23. Balakrishna S, Thirumaran M, Solanki VK (2020) IoT sensor data integration in healthcare using semantics and machine learning approaches. *A handbook of internet of things in biomedical and cyber-physical system*. Springer, Cham, pp 275–300
24. Vanani IR, Amirhosseini M (2021) IoT-based diseases prediction and diagnosis system for healthcare. *The internet of things for healthcare technologies*. Springer, Singapore, pp 21–48
25. Muhammad LJ, Islam MM, Usman SS, Ayon SI (2020) Predictive data mining models for novel coronavirus (COVID-19) infected patients recovery. *SN Comput Sci* 1(4):1–7
26. Kadhim KT, Alsahlany AM, Wadi SM, Kadhum HT (2020) An overview of patients health status monitoring system based on internet of things (IoT). *Wireless Pers Commun* 114(3)
27. Otoom M, Ootoum N, Alzubaidi MA, Etoom Y, Banihani R (2020) An IoT-based framework for early identification and monitoring of COVID-19 cases. *Biomed Signal Process Control* 62:102149
28. Kumar K, Kumar N, Shah R (2020) Role of IoT to avoid spreading of COVID-19. *Int J Intell Netw* 1:32–35
29. Vaishya R, Javaid M, Khan IH, Haleem A (2020) Artificial intelligence (AI) applications for COVID-19 pandemic. *Diab Metab Syndr Clin Res Rev* 14(4):337–339
30. Dey L, Chakraborty S, Mukhopadhyay A (2020) Machine learning techniques for sequence-based prediction of viral-host interactions between SARS-CoV-2 and human proteins. *Biomed J* 43(5):438–450

31. Goodman-Meza D, Rudas A, Chiang JN, Adamson PC, Ebinger J, Sun N et al (2020) A machine-learning algorithm to increase COVID-19 inpatient diagnostic capacity. *PLoS ONE* 15(9):e0239474
32. Khalifa NEM, Taha MHN, Manogaran G, Loey M (2020) A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell. *J Nanopart Res* 22(11):1–13
33. Khanday AMUD, Rabani ST, Khan QR, Rouf N, Din MMU (2020) Machine learning-based approaches for detecting COVID-19 using clinical text data. *Int J Inf Technol* 12(3):731–739
34. Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U et al (2020) Covid-19 outbreak prediction with machine learning. *Algorithms* 13(10):249
35. Punn NS, Sonbhadra SK, Agarwal S (2020) COVID-19 epidemic analysis using machine learning and deep learning algorithms. *MedRxiv*
36. Pinter G, Felde I, Mosavi A, Ghamisi P, Gloaguen R (2020) COVID-19 pandemic prediction for Hungary; a hybrid machine learning approach. *Mathematics* 8(6):890
37. Keshavarzi Arshadi A, Webb J, Salem M, Cruz E, Calad-Thomson S, Ghadirian N et al (2020) Artificial intelligence for COVID-19 drug discovery and vaccine development. *Front Artif Intell* 3:65
38. Subudhi S, Verma A, Patel AB (2020) Prognostic machine learning models for COVID-19 to facilitate decision-making. *Int J Clin Pract* 74(12):e13685
39. Li WT, Ma J, Shende N, Castaneda G, Chakladar J, Tsai JC et al (2020) Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC Med Inform Dec Making* 20(1):1–13
40. Zame WR, Bica I, Shen C, Curth A, Lee HS, Bailey S et al (2020) Machine learning for clinical trials in the era of COVID-19. *Stat Biopharm Res* 12(4):506–517
41. Mick E, Kamm J, Pisco AO, Ratnasiri K, Babik JM, Calfee CS, et al (2020) Upper airway gene expression differentiates COVID-19 from other acute respiratory illnesses and reveals suppression of innate immune responses by SARS-CoV-2. *medRxiv*
42. Karaboga D, Basturk B (2008) On the performance of artificial bee colony (ABC) algorithm. *Appl Soft Comput* 8(1):687–697
43. Moosa JM, Shakur R, Kaykobad M, Rahman MS (2016) Gene selection for cancer classification with the help of bees. *BMC Med Genom* 9(2):135–165
44. Islam MM, Iqbal H, Haque MR, Hasan MK (2017) Prediction of breast cancer using support vector machine and K-Nearest neighbors. In: *Proceedings of the 2017 IEEE region 10 humanitarian technology conference (R10-HTC)*. IEEE, pp 226–229
45. Çayır A, Yenidoğan I, Dağ H (2018) Feature extraction is based on deep learning for some traditional machine learning methods. In: *Proceedings of the 2018 3rd international conference on computer science and engineering (UBMK)*. IEEE, pp 494–497
46. Arowolo MO, Adebisi MO, Adebisi AA, Okesola OJ (2020) A hybrid heuristic dimensionality reduction methods for classifying malaria vector gene expression data. *IEEE Access* 8:182422–182430
47. Okolie PC, Nwadike EC, Chukwunke JL, Nnodim CT (2017) Analysis of cigarette production using double exponential smoothing model. *Acad J Sci* 7(2):293–308
48. Arowolo MO, Isiaka RM, Abdulsalam SO, Saheed YK, Gbolagade KA (2017) A comparative analysis of feature extraction methods for classifying colon cancer microarray data. *ICST Trans Scal Inf Syst* 4(14):153147
49. Arowolo MO, Adebisi M, Adebisi AA, Okesola JO (2020) PCA model for RNA-Seq malaria vector data classification using KNN and decision tree algorithm

## Authors and Affiliations

**Micheal Olaolu Arowolo<sup>1</sup> · Roseline Oluwaseun Ogundokun<sup>2</sup> · Sanjay Misra<sup>3</sup> · Blessing Dorothy Agboola<sup>4</sup> · Brij Gupta<sup>5,6,7</sup>**

✉ Brij Gupta  
bbgupta@nitkkr.ac.in; gupta.brij@gmail.com

Micheal Olaolu Arowolo  
Arowolo.olaolu@lmu.edu.ng

Roseline Oluwaseun Ogundokun  
rosogu@ktu.lt; ogundokun.roseline@lmu.edu.ng

Sanjay Misra  
ssopam@gmail.com

Blessing Dorothy Agboola  
agboola.blessing@lmu.edu.ng

- <sup>1</sup> Department of Computer Science, Landmark University, Omu-Aran, Nigeria
- <sup>2</sup> Department of Multimedia Engineering, Kaunas University of Technology, Kaunas, Lithuania
- <sup>3</sup> Department of Computer Science and Communication, Ostfold University College, Halden, Norway
- <sup>4</sup> Department of Civil Engineering, Landmark University, Omu-Aran, Nigeria
- <sup>5</sup> Department of Computer Science and Information Engineering, Asia University, Taichung 40704, Taiwan
- <sup>6</sup> King Abdulaziz University, Jeddah 21589, Saudi Arabia
- <sup>7</sup> National Institute of Technology Kurukshetra, Kurukshetra 136119, Haryana, India