International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 4, August 2023, pp. 4605~4611
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i4.pp4605-4611

Efficient machine learning classifier to detect and monitor COVID-19 cases based on internet of things framework

Felcia Bel, Sabeen Selvaraj
Department of Computer Science, SRM Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received Jun 2, 2022
Revised Dec 1, 2022
Accepted Dec 7, 2022

Keywords:

Coronavirus disease
Detection and monitoring
Internet of things
Machine learning algorithms
Treatment history

ABSTRACT

In this research work, coronavirus disease 2019 (COVID-19) has been considered to help mankind survive the present-day pandemic. This research is helpful to monitor the patients newly infected by the virus, and patients who have already recovered from the disease, and also to study the flow of virus from similar health issues. In this paper, an internet of things (IoT) framework has been developed for the early detection of suspected cases. This framework is used for collecting and uploading symptoms (data) through sensor devices to the physician, data analytics center, cloud, and isolation/health centers. The symptoms of the first wave, second wave, and omicron are used to identify the suspects. Five machine learning algorithms which are considered to be the best in the existing literature have been used to find the best machine learning classifier in this research work. The proposed framework is used for the rapid detection of COVID-19 cases from real-world COVID-19 symptoms to mitigate the spread in society. This model also monitors the affected patient who has undergone treatment and recovered. It also collects data for analysis to perform further improvements in algorithms based on daily updated information from patients to provide better solutions to mankind.

Corresponding Author:

Sabeen Selvaraj
Department of Computer Science, SRM Institute of Science and Technology
Kattankulathur, Chennai-603203, India
Email: sabeens@srmist.edu.in

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a life-threatening disease for humanity around the globe. It was a great deal for people to survive in this pandemic era. Many variants are arising from this virus frequently. Hence, covid becomes endless despite vaccinations and booster doses. As on January 18, 2022, 328,532,929 confirmed cases and 5,542,359 deaths were reported to World Health Organization (WHO) dashboard [1]. The symptoms due to alpha, delta, and gamma (delta plus) and other virus attacks are taken into consideration for the analysis to find a better internet of things (IoT) solution to mitigate the spread of diseases from future variants. The spread could also be reduced by wearing masks, social distancing, and sanitizing the hands frequently. The government has insisted that people get vaccinated to reduce the effect of this disease.

The research [2], [3] are trying to find a better solution to mitigate the spread of disease rapidly based on early detection and observing new cases. The data which are collected through m-health, telehealth, and real-time patient status could be monitored [4]. A model in [5] is used to analyze the potential cases, confirmed cases, treatment given to confirmed cases, and relevant information about the virus nature for further study. IoT services provided in healthcare were previously used in existing systems [6]. Smart healthcare systems incorporate the idea of implementing health sensors and cloud technology along with IoT.

These smart sensors connected to the human body are used to send and receive data actively [7]. Recently, 5G technologies are also used to connect healthcare environments using IoT [8]. In urban areas, the IoT devices connected to healthcare centers face problems such as security, and privacy. So, the IoT data is processed using a vehicular ad-hoc network (VANET) zone and evaluated using simulators [9]. To acquire patient information, wearable devices are being used. In healthcare, IoT and cloud technology were used in existing systems [10]. IoT-based approaches are also used to find corona cases using fuzzy inference systems [10]. Biometric-based authentication technology is used to secure IoT-based health data [12].

This research work considers data from real-time COVID-19 symptoms using wearable sensor devices. For the early detection of potential cases, the five best machine learning algorithms from the literature are considered in this research work. This research is also used to find the accurate machine learning classifier for the detection of COVID-19 cases for early diagnosis and spread.

2. METHOD

2.1. IoT environment

Wearable sensors and other devices such as wireless devices communicate among themselves to enhance the physical things to smart things [13]. This will improve the quality of human lives. The IoT architecture is divided into three layers [14]. The physical things are connected to sensors to collect heterogeneous data. Collecting more data can also be used to take better decisions. The network layer uploads the collected data and also provides high privacy and security for IoT devices [15]. Sensor data processing is vital in IoT devices [16]. Machine learning or deep learning techniques are used to take better decisions when dealing with massive data [17]. Efficient classification of machine learning or deep learning techniques is used for IoT security [18]. The design carried out by the internet of medical things (IoMT) [19] is highly used to secure data to provide privacy among patients' data. While using IoMT, there are cyber-attacks on patients' information which might become a life-threatening issue for the patient [20]. To overcome this issue, machine learning or deep learning methods are used.

2.2. Proposed architecture

The proposed architecture deals with the early identification and monitoring of COVID-19 suspected cases. The architecture comprises five components. They are i) COVID-19 real-time symptoms collection through biosensors, ii) IoT cloud technology, iii) data analysis center, iv) physician/health care providers, and v) quarantine/isolation/hospital buildings. The data could be analyzed using five machine learning algorithms support vector machine (SVM), neural network (NN), naïve Bayes (NB), k-nearest neighbor (KNN), and decision table (DT) to find the best model for the proposed work. Figure 1 explains the overall IoT architecture of the proposed work.

2.2.1. Dataset

The real-time symptom data are collected using wearable biosensors [21]. Various biosensors are used to collect different symptoms such as fever, cough, sore throat, fatigue, SpO2, shortness of breath. Fever is identified using temperature-based sensors [22]. Audio-based sensors with acoustic and aerodynamic models are used to detect cough [23]. Fatigue is detected using motion-based and heart-rate sensors [24]. Image-based classification is used to detect sore throats [25]. Transmissive pulse oximetry [26] is used to detect SpO2. Oxygen-based sensors are used to detect shortness of breath [27]. Diarrhea, travel, and contact history of patients are collected for 28 days from the day of occurring symptoms using mobile applications.

2.2.2. IoT cloud

The IoT cloud stores the symptom data collected through sensors. The patient information is stored in the cloud through the internet and mobile application. This data will be accessed by data analysts for prediction and the health providers access the patient information for giving recommendations and suggestions. The hospitals/quarantine/isolation centers will also store the inpatient's updated information in the cloud for further study. This helps in a better understanding of the virus, which will be useful in finding suspected cases in a better way to protect people from the spread of disease. This data could be analyzed in the future to tune the system, based on updated routine information from health centers for an efficient diagnosis and treatment of affected patients.

2.2.3. Data analysis center

The patient's symptom information and history are collected from the cloud. The algorithm is trained based on the data received to build the best COVID-19 classification machine learning model which is a key for a physician for monitoring patients from society rapidly. This developed model is useful for humans to take precautions in advance to safeguard themselves from the virus.

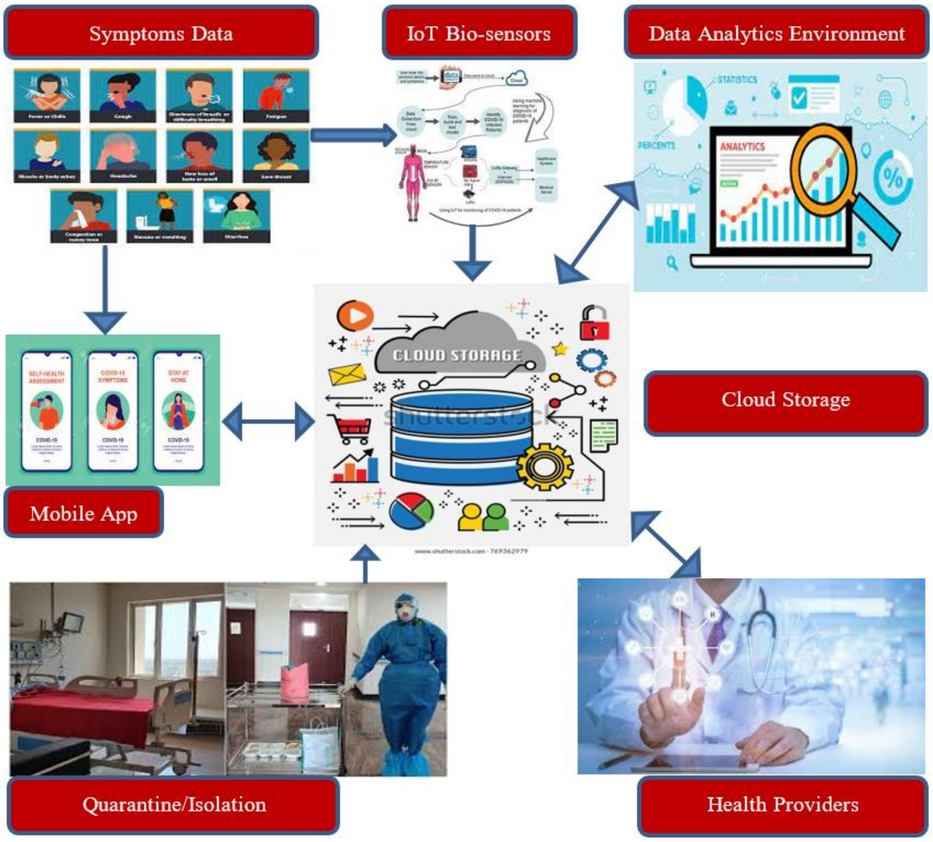

Figure 1. Overall IoT architecture of proposed work

2.2.4. Health providers/physicians

Physicians can monitor the uploaded symptoms data to find suspected cases of COVID-19 patients using the proposed machine learning model. They give proper clinical investigations to reduce the spread of disease. Also, it will allow the confirmed cases to isolate and get treated earlier. Physicians can further connect with patients who were recovered from the disease. To learn updates about their progress and to get knowledge about virus influence for a specific period shall be studied for future enhancement in the medical field.

2.2.5. Quarantine/isolation/hospitals

The records of data of patients were collected directly from quarantine/isolation/hospitals. The patient's history such as age, weight, height, previous histories, contact, travel histories, and other relevant information has also been collected. The treatment chart has been recorded in a cloud of each COVID-19 case.

2.3. Analyzing potential cases

The data analysis center component of the proposed work deals with five machine learning algorithms for early detection and monitoring of COVID-19 cases using IoT. The real-time symptom dataset which contains nine attributes with 1,473 tuples has been taken. In particular, only symptom data has been considered along with travel and contact information of the suspected cases. The result and discussions are discussed in further sections.

2.3.1. Data pre-processing

The dataset with 1,500 tuples was collected from the CORD-19 repository, out of which 1,473 tuples have been taken after preprocessing. From the first and second-wave symptoms, nine major symptoms have been considered for this research work. According to Otoom *et al*. [5], only the first-wave symptoms along with travel and contact information were considered.

2.3.2. Research work

The objective of this research is to predict, whether the suspected person is infected with COVID-19 or not. To do this research, SVM, neural networks, NB, KNN, and DT machine learning classifiers are selected and compared. This work has been done to find the best accuracy among the five classifiers applied in the proposed model. SVM is one of the supervised learning algorithms. From the dataset, consider the training set in which each tuple belongs to either a positive or negative class. This SVM algorithm learns the hyperplane from which the tuples are segregated as positive and negative classes, and how the datapoints and hyperplane margins are maximized. Thus, this trained hyperplane is used to predict a class label for a newly given test tuple. Artificial neural network (ANN) is also another supervised learning method. The machine learns the process exactly like how our human brain thinks. In order to learn the process, multiple layers of nodes are connected through edges. The edges are assigned numerical weights. All inputs to a particular node are computed to produce an output. The training sets of data are given to ANN to learn the weight assigned to best classify the tuple for each class. This learnt model is used to predict a class label for a test tuple. NB has supervised learning which is based on a probabilistic approach. Bayes theorem is used for computing the parameters. The training sets are given to the model to compute the class label based on the multiple model parameters. The test tuple class label has been predicted by learning the process from the training set. KNN is a supervised learning method that follows a lazy approach. Training set tuples are used to compute the model for class labels. The distance between the training set and the test set will be used to compute the model for prediction. This is done by grouping the class labels of both training tuples along with test tuples. The DT is also a supervised learning method. The method learns the process by the given training set with class labels applying input conditions. It builds a DT that is used to compute test sets for better prediction. The table comprises a set of rules and appropriate actions.

2.4. Performance evaluation

The performance and model evaluation measures are applied in five machine learning algorithms. The measures are accuracy, root mean square error (RMSE), precision, recall, F-measure, sensitivity, specificity and geometric mean. The accuracy of the classifier is computed based on the best performance measures.

2.4.1. Confusion matrix

This is to envisage the supervised learning problem into 2×2 matrix in which each row represents the actual class, and each column represents the predicted class. It consists of four values, which are true positive, false positive, false negative, and true negative. The representation of the confusion matrix was given in Table 1.

Table 1. Confusion matrix representation

		Predicted class	
		P	N
Actual class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

2.4.2. Cross validation and geometric mean

This is one of the machine learning techniques used for finding accuracy in randomly sampled partitions for given tuples. By using this technique, the overall computation time of the model becomes less. It will also provide timely predictions. This is carried out by 10 folds in which, data is segregated as nine folds which are used for training and one-fold is used for testing. Accuracy, precision, recall, F-measure, and RMSE are computed using (1) to (4). Sensitivity (SNR) in (5) refers to true positive rate and specificity (SPC) in (6) refers to false positive rate (FPR) which is used to calculate the geometric mean as in (7) known as model evaluation metrics.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{MSE} = \sqrt{\frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}} \tag{2}$$

$$\text{F measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

where

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}}, \text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (4)$$

$$\text{SNS} = \frac{\text{TP}}{\text{TP + FN}} \quad (5)$$

$$\text{SPC} = \frac{\text{FP}}{\text{FP + TN}} \quad (6)$$

$$\text{Geometric Mean, GM} = \sqrt{\text{SNS} - \text{SPC}} \quad (7)$$

Figure 2 shows the confusion matrix for all five classifiers in which 10-fold cross-validation has been applied. The left upper and right lower values are predicted correctly and the right upper and left lower values are wrong predictions. The information given below is a detailed description of Figure 2.

- In SVM, as shown in Figure 2(a), 621 tuples are actually positive, the classifier also predicted as positive and 749 tuples are actually negative, the classifier also predicted as negative. Actually, 103 tuples are negative, but the classifier predicted as positive, and 0 tuple are actually positive, but the classifier predicted as negative.
- In NN, as shown in Figure 2(b), 620 tuples are actually positive, the classifier also predicted as positive and 749 tuples are actually negative, the classifier also predicted as negative. Actually, 103 tuples are negative, but the classifier predicted as positive, and 1 tuple are actually positive, but the classifier predicted as negative.
- In NB, as shown in Figure 2(c), 585 tuples are actually positive, the classifier also predicted as positive and 749 tuples are actually negative, the classifier also predicted as negative. Actually, 103 tuples are negative, but the classifier predicted as positive and 36 tuples are actually positive, but the classifier predicted as negative.
- In KNN, as shown in Figure 2(d), 621 tuples are actually positive, the classifier also predicted as positive and 748 tuples are actually negative, the classifier also predicted as negative. Actually, 104 tuples are negative, but the classifier predicted as positive, and 0 tuple are actually positive, but the classifier predicted as negative.
- In DT, as shown in Figure 2(e), 621 tuples are actually positive, the classifier also predicted as positive and 749 tuples are actually negative, the classifier also predicted as negative. Actually, 103 tuples are negative, but the classifier predicted as positive, and 0 tuple are actually positive, but the classifier predicted as negative.

621	0
103	749

(a)

620	1
103	749

(b)

585	36
103	749

(c)

621	0
104	748

(d)

621	0
103	749

(e)

Figure 2. Confusion matrices (a) SVM, (b) NN, (c) NB, (d) KNN, and (e) DT

3. RESULTS AND DISCUSSION

Table 2 summarizes the performance measures of all the five classifiers calculated for accuracy, RMSE, F-measure, and geometric mean in percentage using the corresponding (1) to (7). Based on the results obtained, it is concluded that the DT acquires the highest accuracy to predict the suspected cases based on the IoT framework. Results obtained in Table 2 are graphically represented in Figure 3. The performance measures namely accuracy, F-measure, RMSE, and geometric mean are shown in Figures 3(a) to 3(d) respectively.

Table 2. Results of all five algorithms based on performance

Classifiers	Accuracy	F-measure	Root mean square	Geometric mean
SVM	93.01%	92.34%	26.44%	34.77%
NN	92.94%	92.33%	26.57%	34.54%
NB	90.56%	89.38%	30.72%	25.08%
KNN	92.94%	92.28%	26.57%	34.94%
DT	93.55%	93.02%	25.396%	33.63%

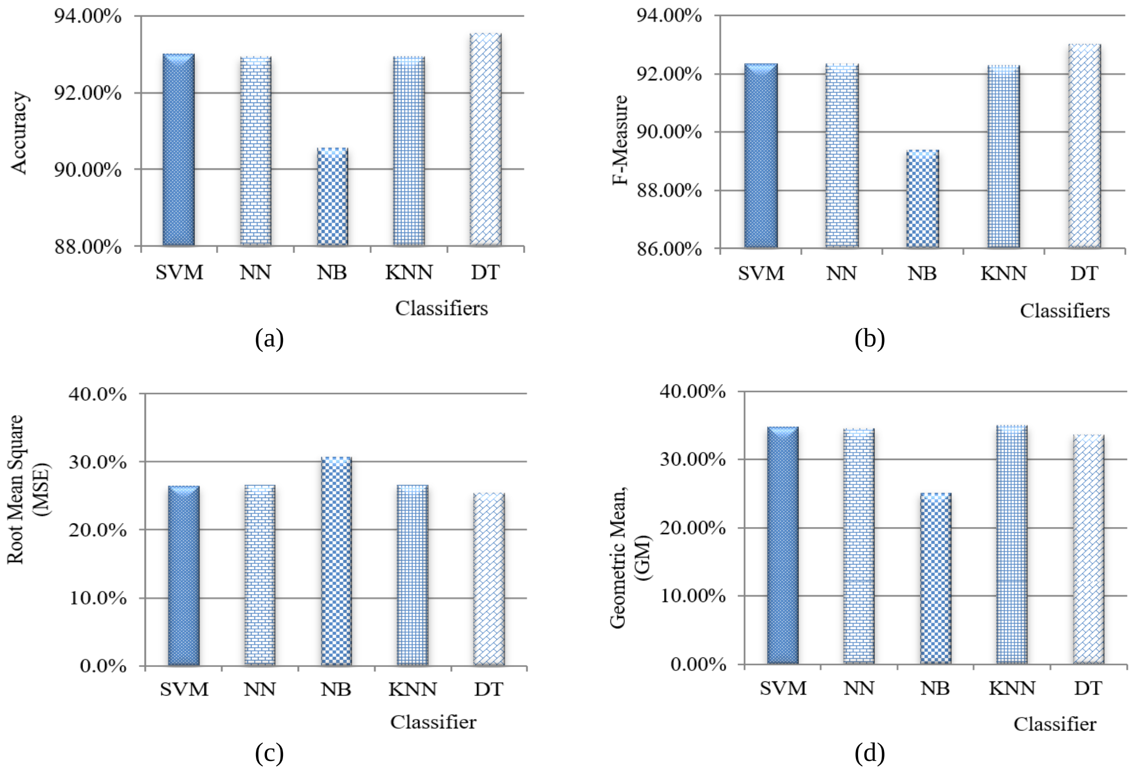

Figure 3. Results of performance measures obtained for five classifiers (a) accuracy, (b) F-measure, (c) RMSE, and (d) geometric mean

4. CONCLUSION

An IoT framework has been developed in this paper to identify the covid cases from the data provided by the health care centers. The previous history of COVID-19 patients' information has been taken by the analysts from the cloud to develop a machine learning model for predicting the cases rapidly. This model helps people to isolate/quarantine from other people in society. The proposed work will reduce the spread of disease and will also alert people to be aware by giving proper care. Furthermore, it will be used as a base model for future variants, which are evolving now from coronavirus. The proposed work also monitors the confirmed patient records for further study and treatment. The proposed model has been experimented with using five machine learning algorithms like SVM, neural network, NB, KNN, and DT in which, the decision table has been selected as the best model for IoT-based early detection. By applying performance measures in each algorithm mentioned above, the decision tree acquires 93.33% accuracy to detect and monitor COVID-19 cases using the IoT framework.

REFERENCES

[1] WHO, "WHO COVID-19 Dashboard," *World Health Organization*, 2022. https://covid19.who.int/ (accessed Jan. 18, 2022).
[2] M. P. Kelly, "Digital technologies and disease prevention," *American Journal of Preventive Medicine*, vol. 51, no. 5, pp. 861–863, 2016, doi: 10.1016/j.amepre.2016.06.012.
[3] P. M. Hlaing, T. R. Nopparatjamjomras, and S. Nopparatjamjomras, "Digital technology for preventative health care in Myanmar," *Digital Medicine*, vol. 4, no. 3, 2018, doi: 10.4103/digm.digm.
[4] B. Dinesen *et al.*, "Personalized telehealth in the future: A global research agenda," *Journal of Medical Internet Research*, vol. 18, no. 3, 2016, doi: 10.2196/jmir.5257.
[5] M. Otoom, N. Otoum, M. A. Alzubaidi, Y. Etoom, and R. Banihani, "An IoT-based framework for early identification and monitoring of COVID-19 cases," *Biomedical Signal Processing and Control*, vol. 62, Sep. 2020, doi: 10.1016/j.bspc.2020.102149.
[6] M. Usak, M. Kubiatko, M. S. Shabbir, O. V. Dudnik, K. Jermsittiparsert, and L. Rajabion, "Health care service delivery based on the internet of things: a systematic and comprehensive study," *International Journal of Communication Systems*, vol. 33, no. 2, pp. 1–17, 2020, doi: 10.1002/dac.4179.
[7] N. Nasser, Q. Emad-ul-Haq, M. Imran, A. Ali, I. Razzak, and A. Al-Helali, "A smart healthcare framework for detection and monitoring of COVID-19 using IoT and cloud computing," *Neural Computing and Applications*, vol. 7, 2021, doi: 10.1007/s00521-021-06396-7.
[8] M. Chen, J. Yang, Y. Hao, S. Mao, and K. Hwang, "A 5G cognitive system for healthcare," *Big Data and Cognitive Computing*, vol. 1, no. 1, pp. 1–15, 2017, doi: 10.3390/bdcc1010002.

[9] M. Rath and B. Pattanayak, "Technological improvement in modern health care applications using internet of things (IoT) and proposal of novel health care approach," *International Journal of Human Rights in Healthcare*, vol. 12, no. 2, pp. 148–162, 2019, doi: 10.1108/IJHRH-01-2018-0007.

[10] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of internet of things and cloud computing on healthcare systems: opportunities, challenges, and open problems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4151–4166, 2019, doi: 10.1007/s12652-017-0659-1.

[11] N. Anjum et al., "IoT-based COVID-19 diagnosing and monitoring systems: a survey," *IEEE Access*, vol. 10, pp. 87168–87181, 2022, doi: 10.1109/ACCESS.2022.3197164.

[12] H. Hamidi, "An approach to develop the smart health using internet of things and authentication based on biometric technology," *Future Generation Computer Systems*, vol. 91, pp. 434–449, 2019, doi: 10.1016/j.future.2018.09.024.

[13] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: a survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015, doi: 10.1109/COMST.2015.2444095.

[14] P. Sethi and S. R. Sarangi, "Internet of things: architectures, protocols, and applications," *Journal of Electrical and Computer Engineering*, pp. 1–25, 2017, doi: 10.1155/2017/9324035.

[15] G. Yang, "An overview of current solutions for privacy in the internet of things," *Frontiers in Artificial Intelligence*, vol. 5, no. March, 2022, doi: 10.3389/frai.2022.812732.

[16] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 414–454, 2014, doi: 10.1109/SURV.2013.042313.00197.

[17] E. Ahmed *et al.*, "The role of big data analytics in internet of things," *Computer Networks*, vol. 129, pp. 459–471, 2017, doi: 10.1016/j.comnet.2017.06.013.

[18] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," *IEEE Communications Surveys and Tutorials*, vol. 22, no. 3, pp. 1646–1685, 2020, doi: 10.1109/COMST.2020.2988293.

[19] C. Camara, P. Peris-Lopez, and J. E. Tapiador, "Security and privacy issues in implantable medical devices: A comprehensive survey," *Journal of Biomedical Informatics*, vol. 55, pp. 272–289, 2015, doi: 10.1016/j.jbi.2015.04.007.

[20] R. Altawy and A. M. Youssef, "Security tradeoffs in cyber physical systems: a case study survey on implantable medical devices," *IEEE Access*, vol. 4, no. c, pp. 959–979, 2016, doi: 10.1109/ACCESS.2016.2521727.

[21] M. A. Alzubaidi, M. Otoom, N. Otoum, Y. Etoom, and R. Banihani, "A novel computational method for assigning weights of importance to symptoms of COVID-19 patients," *Artificial Intelligence in Medicine*, vol. 112, no. September 2020, Art. no. 102018, 2021, doi: 10.1016/j.artmed.2021.102018.

[22] J. Medina, M. Espinilla, L. García-Fernández, and L. Martínez, "Intelligent multi-dose medication controller for fever: From wearable devices to remote dispensers," *Computers and Electrical Engineering*, vol. 65, pp. 400–412, 2018, doi: 10.1016/j.compeleceng.2017.03.012.

[23] Y. Umayahara, Z. Soh, K. Sekikawa, T. Kawae, A. Otsuka, and T. Tsuji, "A mobile cough strength evaluation device using cough sounds," *Sensors (Switzerland)*, vol. 18, no. 11, 2018, doi: 10.3390/s18113810.

[24] Q. Abbas and A. Alsheddy, "Driver fatigue detection systems using multi-sensors, smartphone, and cloud-based computing platforms: A comparative analysis," *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–38, 2021, doi: 10.3390/s21010056.

[25] B. Askarian, S. C. Yoo, and J. W. Chong, "Novel image processing method for detecting strep throat (Streptococcal pharyngitis) using smartphone," *Sensors (Switzerland)*, vol. 19, no. 15, pp. 1–17, 2019, doi: 10.3390/s19153307.

[26] M. Nitzan, A. Romem, and R. Koppel, "Pulse oximetry: Fundamentals and technology update," *Medical Devices: Evidence and Research*, vol. 7, no. 1, pp. 231–239, 2014, doi: 10.2147/MDER.S47319.

[27] A. Gaidhani, K. S. Moon, Y. Ozturk, S. Q. Lee, and W. Youm, "Extraction and analysis of respiratory motion using wearable inertial sensor system during trunk motion," *Sensors (Switzerland)*, vol. 17, no. 12, 2017, doi: 10.3390/s17122932.

BIOGRAPHIES OF AUTHORS

Felcia Bel received a B.Sc. degree in physics from Manonmanium Sundranar University and M.C.A. degree in computer applications from Bharadhidasan University, in 2002 and 2005, respectively, and M.Phil. degree in computer science in 2008. She had been working as an assistant professor in various colleges from 2008 to 2020. At present, she is pursuing research at SRM Institute of Science and Technology, Chennai 603203, Tamil Nadu, India. Her research interests are the Internet of things, data mining, and machine learning. She can be contacted at fh4219@srmist.edu.in.

Sabeen Selvaraj is working at the Department of Computer Science, SRM Institute of Science and Technology. He received his doctoral degree from Anna University in 2012. He received his postgraduate degree MCA in the year 2002 and M.Tech (CSE) in the year 2015. He has 20 years of teaching experience in various engineering colleges: Mohamad Sathak Engineering College, Noorul Islam College of Engineering, Jaya Engineering College and Sidharth Institute of Engineering and Technology, India. He is a life member of ISTE. His area of interests includes data mining, IoT, machine learning. He has published many papers in international journals. He has guided more than 100 projects for the post graduate students. He can be contacted at sabeens@srmist.edu.in.