

Received February 16, 2021, accepted March 7, 2021, date of publication March 17, 2021, date of current version March 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3066365

# Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems

ROMANY FOUAD MANSOUR<sup>1</sup>, ADNEN EL AMRAOUI<sup>2</sup>, ISSAM NOUAOURI<sup>2</sup>, VICENTE GARCÍA DÍAZ<sup>3</sup>, DEEPAK GUPTA<sup>4</sup>, (Senior Member, IEEE), AND SACHIN KUMAR<sup>5</sup>, (Senior Member, IEEE)

<sup>1</sup>Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

<sup>2</sup>Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A), Université d'Artois, 62400 Béthune, France

<sup>3</sup>Department of Computer Science, University of Oviedo, 33007 Oviedo, Spain

<sup>4</sup>Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, Delhi 110086, India

<sup>5</sup>Department of Computer Science, South Ural State University, 454080 Chelyabinsk, Russia

Corresponding author: Sachin Kumar (sachinagnihotri16@gmail.com)

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Government Order) under Grant FENU-2020-0022.

**ABSTRACT** The recent advancements in Internet of Things (IoT), cloud computing, and Artificial Intelligence (AI) transformed the conventional healthcare system into smart healthcare. By incorporating key technologies such as IoT and AI, medical services can be improved. The convergence of IoT and AI offers different opportunities in healthcare sector. In this view, the current research article presents a new AI and IoT convergence-based disease diagnosis model for smart healthcare system. The major goal of this article is to design a disease diagnosis model for heart disease and diabetes using AI and IoT convergence techniques. The presented model encompasses different stages namely, data acquisition, preprocessing, classification, and parameter tuning. IoT devices such as wearables and sensors permit seamless data collection while AI techniques utilize the data in disease diagnosis. The proposed method uses Crow Search Optimization algorithm-based Cascaded Long Short Term Memory (CSO-CLSTM) model for disease diagnosis. In order to achieve better classification of the medical data, CSO is applied to tune both 'weights' and 'bias' parameters of CLSTM model. Besides, isolation Forest (iForest) technique is employed in this research work to remove the outliers. The application of CSO helps in considerable improvement in the diagnostic outcomes of CLSTM model. The performance of CSO-LSTM model was validated using healthcare data. During the experimentation, the presented CSO-LSTM model accomplished the maximum accuracies of 96.16% and 97.26% in diagnosing heart disease and diabetes respectively. Therefore, the proposed CSO-LSTM model can be employed as an appropriate disease diagnosis tool for smart healthcare systems.

**INDEX TERMS** Internet of Things, convergence, cloud computing, artificial intelligence, smart healthcare, disease diagnosis.

## I. INTRODUCTION

Healthcare sector started leveraging information technology in the recent years to develop modern applications and enhance the diagnostic and treatment processes. Advanced techniques and scientific theory are the major entities that generate huge volumes of digital data. Followed by, advanced clinical applications are the brainchildren of information technology which have been developed in the recent times.

The associate editor coordinating the review of this manuscript and approving it for publication was Abdulsalam Yassine<sup>1</sup>.

Also, advanced healthcare is assumed to have simple, elegant, and multi-tasking applications. These modifications are incorporated as clinical model extension (from disease-based to patient-based care), informatization development changes (from medical data to regional medical data), extension in clinical management (general management to personal management), and modifications from prevention and treatment (Shifting of concentration from disease treatment to preventive medical system) [1]. Hence, the following changes are concentrated to satisfy the basic requirements of individuals in order to enhance the proficiency of health care which in

turn improve health service knowledge and imply the future deployment of smart medicine.

Advanced medical services are comprised of several stakeholders such as doctors, patients, clinical and research centers. Multiple dimensions should be considered such as disease preventive measures and observation, prognosis and treatment, clinical management, health decision-making, and medical studies. For instance, mobile internet, Cloud Computing (CC), big data, 5G systems, microelectronics, and Artificial Intelligence (AI), along with smart biotechnology are assumed to be the milestones of modern healthcare. These methodologies are employed in every stage of advanced healthcare. From patients' viewpoint, wearable or portable devices can be applied for monitoring their health condition whenever required. They can look for clinical guidance through virtual support and control their homes remotely using remote facilities. From the perception of doctors, smart clinical decision support systems can be applied to guide and enhance the diagnostic procedures.

An extensive diffusion and deployment of effectively-incorporated hardware and modern medical sensors for unique healthcare has intended to develop a new concept called Internet of Medical Things (IoMT). It modifies the healthcare process and number of medical devices that use IoT to achieve better profit in future [2]. The data, captured with the help of portable, ingestible, and incorporated sensors, mobile patterns and device usage patterns, enable the researcher to track a user's habits. With further data collection, it is possible to reveal their medical status by applying state-of-the-art as well as Machine Learning (ML) or Deep Learning (DL)-based methods. Classical cloud technology that relies on structures for big data analysis, is applied to provide optimal performance, scalability and support non-safety as well as delay-based IoT domains. However, if a patient is serious with the availability of limited resources and when they require high degree of efficiency and accessibility, the disconnection from main network or latency difference might dramatically produce negative effect and results in dreadful consequences under emergency cases. The rapid development of structures that examine the collaboration of cloud, fog, and edge computing remains a challenging process yet. The main aim of this method is to apply complete edge nodes and low-level fog nodes to manage functional tasks with regards to data processing, examination, correlation, and inference. Hence, the above-mentioned approaches produce challenging outcomes by implementing scalable medical domain services. This occurs because the smart mapping of processing and resource management operations overcome the nodes to satisfy the fundamental needs of IoMT model [3].

Under the application of Artificial Intelligence (AI) models, surgical devices, and mixed reality applications, both diagnosis and disease treatment are highly robust [4], [5]. By using AI, specific outcomes are attained from Clinical Decision Support System (CDSS) such as the diagnosis of hepatitis, lung tumor, and skin cancer. Further, the accuracy of

AI diagnosis has exceeded the accuracy accomplished manually. In addition, ML-based models are precise in comparison with well-trained physicians, particularly pathologists and imaging experts. Therefore, a remarkable and representative product in CDSS was rolled out by IBM's Watson. This product has an effective cognitive mechanism and is used for providing the best solution with the help of in-depth analysis of medical and literature details. As a result, a drastic effect has been experienced by healthcare professionals in diagnosing both diabetes and cancer. The application of CDSS is highly efficient and supports the physicians in enhancing the diagnostic processes, limiting the incidence of unexploited diagnosis as well as misdiagnosis, and enabling the users receive timely and proper medical treatment. According to smart diagnosis, patient's health state and disease severity can be defined accurately to follow a personalized treatment procedure.

The current research work presents a new AI and IoT convergence-based disease diagnosis model for smart healthcare system. The aim is to develop a disease diagnosis model using AI and IoT convergence technique for the diagnosis of diabetes and heart disease. The presented model encompasses different stages namely, data acquisition, preprocessing, classification, and parameter tuning. IoT devices such as wearables and sensors perform the data acquisition process whereas AI techniques process this data to diagnose the disease. The proposed AI and IoT convergence method makes use of Crow search Optimization algorithm-based Cascaded Long Short Term Memory (CSO-CLSTM) model for disease diagnosis. Besides, isolation Forest (iForest) technique is employed in this study to remove the outliers. In order to improve the diagnostic outcome, CSO is applied to tune both 'weights' and 'bias' parameters of CLSTM model. CSO is utilized here since it helps to improve the diagnostic outcome of CLSTM method. The effectiveness of CSO-LSTM model was validated using healthcare data. The contributions of this research article is summarized herewith.

- Designing and development of a novel AI and IoT convergence-based disease diagnosis model for smart healthcare system
- Proposed a CSO-CLSTM model for diagnosing diabetes and heart disease
- Incorporated iForest technique-based outlier detection process to improve the classification results
- Performed parameter tuning of LSTM model using CSO algorithm
- Validated the performance of CSO-LSTM model on two benchmark datasets.

## II. RELATED WORKS

Numerous works have been conducted earlier to develop system that senses the physiological variables and health indicators to assess severe cases and accidents. Initially, Mustlag *et al.* [5] applied Wireless Body Sensor Network (WBSN) to observe the heart rate and movement of

users, whenever they require, even from remote areas. In this study, edge node is connected with internet and it forwards an alert (mobile phone) to family members, whenever important changes occur (early prediction of falls, tachycardia, or bradycardia). In line with this, Villarrubia *et al.* [6] projected a method to monitor the patients from home and their heart rate by computing fundamental examination of electrocardiogram (ECG) information. In the literature [7], an emotion-aware connected healthcare model was developed using an efficient emotion detection module. A set of distinct IoT devices was utilized in this study to capture speech and image signals of a patient in smart homes.

In Kaur and Jasuja [8], the developers examined the application of Bluemix cloud method to record the physiological information and enable remote access by physicians. The simulation outcome is visualized and is processed with respect to IBM Watson IoT environment. In Alwan and Rao [9], a case study was conducted for fever analysis using incorporated system that frequently observes the health data of patients. In Satija *et al.* [10], a real-time IoT based ECG telemetry was projected. In this work, researchers have depicted the efficiency of a model based on diverse activities. Static monitoring reduces the application of domain sensors to collect contextual data and perform multimodal process. Followed by, Pham *et al.* [11] presented a model in which ecological sensors, optitrack cameras, and smartwatch-based sensors are employed to collect video, image, and audio signals with particular wearables for the collection of physiological variables. In literature [12], a novel smart healthcare model was proposed which comprised of a pathology detection technique using deep learning technique. Pathogens can be identified from electroencephalogram signals of a patient. In this model, a smart EEG headset captures the EEG signals and transmits it to a mobile edge computing server. The server pre-processes the signals and sends it to the cloud server.

Uddin [13] projected a solution to examine various human activities with the help of wearable sensors as well as Long Short-term Memory-Recurrent Neural Network (LSTM-RNN) which were implemented on local fog server and GPU acceleration. In the study conducted earlier [14], additional sensors were employed for movement tracking and to examine the application of Support Vector Machines (SVM) and Random Forest (RF) classification method for movement forecast. Some of the recently developed models for conducting physiological data analysis in portable sensors simulate the analysis of edge ML approaches. However, there are issues involved in predicting the abnormalities of physiological variables in terms of edge stream computing structure. In this study, Hierarchical Temporal Memory (HTM) was implemented in a distributed manner. The model was implemented on edge nodes and was used for the inference. Additionally, Queralt *et al.* [15] proposed a fall prediction solution based on LSTM RNN method which is executed at edge level. The performance of Multi Access Edge Computing method was defined along with a case study on Electroencephalography (EEG) information.

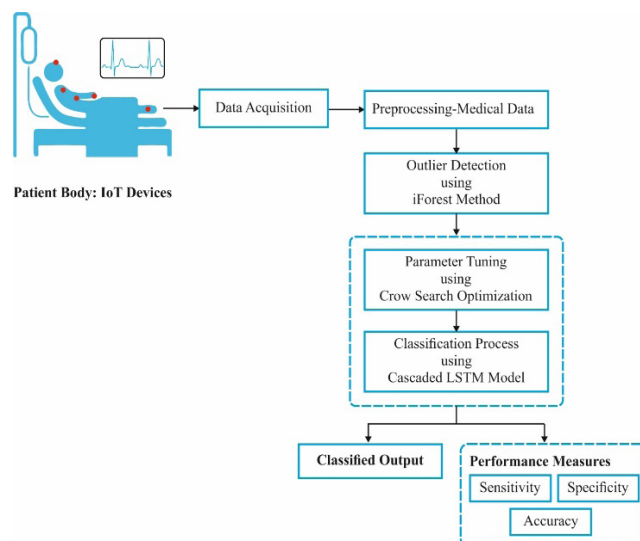


FIGURE 1. Working process of CSO-CLSTM method.

This led to a scenario in which the developers assumed that the major functions are to be executed from edge side and satisfy the application needs (data compression, feature extraction well as classification). The accuracy of the results was compared with existing classification models like RF, Naive Bayes (NB), k-Nearest Neighbors (kNN), and classification or regression trees). Alternatively, the study also used few models to classify anomalies in ECG signals as applied by Azimi *et al.* [16]. Hierarchical Computing Architecture for Healthcare (HiCH) was introduced and its variant i.e., Monitor-Analyze-Plan-Execute Plus Knowledge (MAPE-K) mechanism was implemented by IBM to share the process among three layers called edge, fog, and cloud. In the literature [17], a CNN-based automatic EEG pathology detection model was presented. It used 1D and 2D convolutions to capture temporal and spatial information individually.

### III. THE PROPOSED SMART HEALTHCARE DIAGNOSIS MODEL

Figure 1 shows the overall working process. The proposed approach is effective in terms of former wireless communications and it consumes low power with high freedom of activity for users in external movement. In addition, tiny and low-weight IoT devices are employed in this model which are user-friendly. Some such IoT devices are smartphones, wrist-band, smartwatch, and so forth.

The embedded sensors are employed to proceed with extensive computations to estimate and distinguish normal and abnormal heart rates. The subjects are embedded with smart devices for instance smartphones which can be taken anywhere in pockets. Additionally, embedded ECG as well as temperature sensors are highly recommended to collect data regarding heart parameters of the subject. From this data, results of their common lifestyle can also be determined. When data is received through low-power Bluetooth communication, smartphones process the data and categorize it as

either healthy or unhealthy. The android platform performs the prediction of diabetes and efficient heart rate. Initially, IoT devices gather patient data and preprocess it to transform the data into a compatible format. Pre-processing is composed of few stages such as data transformation, format conversions, and class labeling. Then, iForest technique is employed to get rid of outliers exist in patient data. Followed by, CSO-CLSTM model is employed to classify the data into existence and non-existence of the disease.

#### A. iFOREST-BASED OUTLIER REMOVAL PROCESS

The preprocessed medical data is fed into iForest, a tree-based outlier prediction method with linear time complexity and maximum precision. It can be applied for high-dimensional and huge volumes of data. Since the abnormalities are 'low and varied', it is heavily prone to isolation. In case of data-based random tree, the records are cropped until isolation is performed. Random division tends to make outlier short length alike records with distinguishable values. Here, it is recommended to divide in earlier division [18]. iForest is composed of iTrees (Isolation Tree). Every iTree is referred to as a binary tree. The steps involved in execution process are given below.

- i. Select few sample points called subsamples from training data and place them in root node of a tree.
- ii. Point the attribute and produce a cutting point 'p' from recent node data. At the same time, cutting point is produced from maximum as well as minimum values of certain parameter in recent node data.
- iii. A hyperplane is emulated from cutting point. While the data space of recent node is classified into two subspaces namely, data which is minimum than 'p' in certain attribute and is placed on left child and the data which is maximum than 'p' and is placed on right child of the present node.
- iv. Follow steps 2 and 3, till the child node reaches a single record.

Once the iTrees are accomplished, training iForest is terminated. Then the testing data is estimated with the help of generated iForest. In case of testing records, a traversal of all iTrees is considered and the height of each record is determined. Followed by, the average height of a record from each tree is determined. When the average height is minimum than the applied threshold, a record is assumed to be an outlier.

#### B. DISEASE DIAGNOSIS MODEL USING CSO-LSTM MODEL

After the removal of outliers in healthcare data, CSO-CLSTM model is applied to perform classification process. RNNs are special standard Artificial Neural Networks (ANNs) using which time series of long range structural values can be developed. A fundamental theme of RNNs is the inclusion of time delay unit as well as feedback connection, where the data from former state is applied in upcoming stage. The structure of RNN is comprised of input layer, otherwise

called as sequence layer, which applies input as a series of vectors  $\{x^{<1>}, \dots, x^{<z>}, \dots, x^{<Z>}\}$  with features for every time step; Here, network proceeds with a series of hidden activations  $\{a^{<1>}, \dots, a^{<z>}, \dots, a^{<Z>}\}$  as well as the resultant vector  $\{\hat{y}^{<1>}, \dots, \hat{y}^{<z>}, \dots, \hat{y}^{<Z>}\}$  for  $Z$  timesteps. A primary activation  $0^{(0)}$  is allocated as a vector of zeros. Then, both activation as well as final prediction at time  $z$  is illustrated as given herewith.

$$a^{<z>} = g \left( W_a \cdot \left[ a^{<z-1>}, x^{<z>} \right] + b_a \right) \quad (1)$$

$$\hat{y}^{<z>} = g \left( W_y \cdot a^{<z>} + b_y \right) \quad (2)$$

Here, the vector available in square brackets are referred to as a vector combination of activation from existing timestep as well as input from recent timestep,  $W_a$  and  $W_y$  denote activation as well as output weight matrices correspondingly. Here,  $b_a$  and  $b_y$  imply activation and output bias terms. Additionally, operator  $g$  signifies a generic activation function. RNN feature is a neuron of hidden layer which activates the existing time step for computing an activation of recent time step. Thus, for RNN, the detection of final outcome at recent time step  $\hat{y}^{<z>}$  is computed with data from input  $x^{<z>}$ . However, using the data from  $x^{<z>}$  to  $x^{<z-1>}$ , activation  $a^{<z>}$  is done at former time step. It is named as unidirectional RNN since it applies data from old sequence inputs to evaluate the prediction at specific time step. Eqs. (1) and (2) imply forward propagation of RNN. During backward propagation, weights and bias are upgraded with the help of optimization method [19]. Hence, it is called Backpropagation Through Time (BPTT). One of the major complexities in training RNN is its diminishing gradient issues. Further, its partial derivatives are smaller in deep layers for maximum time steps. The network parameters, in this case, cannot be changed in consecutive iterations while the learning process gets terminated. The above-mentioned issues are resolved when RNN unit is replaced with gated cell unit named LSTM unit.

LSTM unit shows a modification in remarkable RNN to capture long-term dependencies and it enables to report the problem of diminishing gradient. Therefore, LSTM memory cell is composed of five modules namely, memory cell  $c^{<z>}$ , candidate value  $\tilde{c}^{<z>}$  to replace the memory cell at every timestep and three other gates such as update gate  $\Gamma_u$ , forget gate  $\Gamma_f$  and output gate  $\Gamma_o$ . Memory cell is applied to record specific values for prolonged time during training process. Assume three gates are derived from 0 and 1. Both weight matrix as well as bias term might get upgraded during training process. Finally, forget gate enables the selection of type of data which is thrown away and is represented as shown below.

$$\Gamma_f = \sigma \left( W_f \cdot \left[ a^{<z-1>}, x^{<z>} \right] + b_f \right) \quad (3)$$

Here, the update gate decides whether to replace the memory cell with candidate value which is defined herewith.

$$\Gamma_u = \sigma \left( W_u \cdot \left[ a^{<z-1>}, x^{<z>} \right] + b_u \right) \quad (4)$$



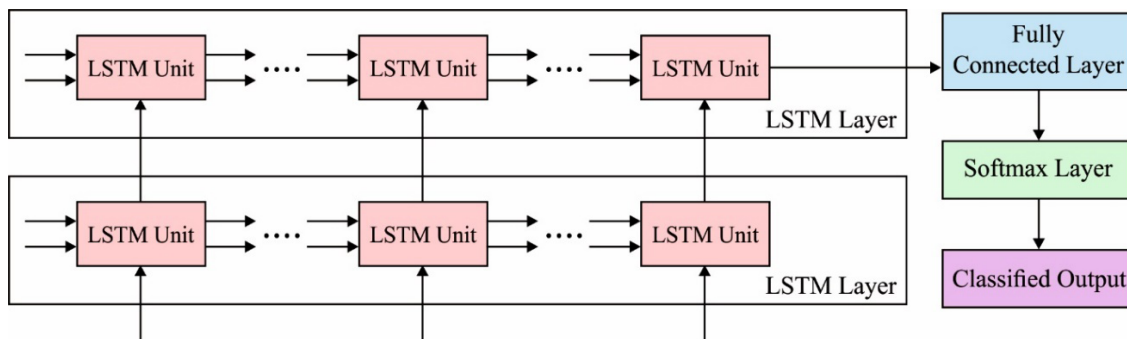


FIGURE 2. The Structure of CLSTM.

Consequently, output gate is a section in which activation at recent time step is produced and demonstrated as follows.

$$\Gamma_0 = \sigma \left( W_0 \cdot [a^{<z-1>}, x^{<z>}] + b_0 \right) \quad (5)$$

In former function,  $\sigma$  implies the sigmoid function. Hence, the function which is used for monitoring the nature of LSTM unit is depicted as follows:

$$\tilde{C}^{<z>} = \tanh \left( W_c \cdot [a^{<z-1>}, x^{<z>}] + b_c \right) \quad (6)$$

$$\tilde{C}^{<t>} = \Gamma_u * \tilde{C}^{<z>} + \Gamma_f * \tilde{C}^{<z-1>} \quad (7)$$

$$a^{<z>} = \Gamma_0 * \tanh \left( C^{<z>} \right) \quad (8)$$

where  $W_c$  and  $b_c$  correspond to cell weight matrix and bias term respectively. The function  $*$  implies Hadamard product and  $\tanh$  denotes hyperbolic tangent function. Figure 2 shows the structure of CLSTM model.

In this approach, a cascade of two RNNs with LSTM units is employed. The primary network applies input as features gained from mRMR model and executes 4-class (W, NI-REM, N2, and N3) classification (N1 and REM epochs are combined within a single class). The alternate network applies the input features estimated by PCA. Followed by, NI-REM epochs and RNN classify the epochs into two classes (namely, NI and REM). Therefore, RNNs are presented in a similar model, where input layer is a sequence layer with 30 time steps; the LSTM layers are applied or else the features from input signals are learnt. Further, Fully Connected (FC) layer is utilized to convert the output size of preceding layers as the number of sleep stages for examination process. Softmax layer is processed as the possibility of target class. The major benefits of utilizing Softmax activation function is that it can generate the output probability range within a limited time. The numerical expression is shown herewith.

$$\hat{y}_j^{(i)} = \frac{e^{z_j^{(i)}}}{\sum_{j=1}^c e^{z_j^{(i)}}} \quad (9)$$

The superscript  $i$  defines a generic training sample, subscript  $j$  implies a generic neuron of FC layer,  $z$  defines the final value of FC layer and  $C$  means the count of target classes. Hence, cost function is reduced in network training

model, where the function of weights,  $W$  and bias term  $b$ , represent the average of cross entropy functions for  $C$ :

$$J(W, b) = -\frac{1}{M} \sum_{i=1}^M \sum_{j=1}^C y_j^{(i)} \cdot \log \left( \hat{y}_j^{(i)} \right) \quad (10)$$

Here,  $M$  defines the count of training sample,  $y$  signifies true label and  $\hat{y}$  depicts the measure evaluated by the system. The measure of  $C$  is allocated as 4 for initial RNN and 2 for secondary RNN. The major difference between these two networks is that the first structure is a single LSTM layer with sequence-to-label manner, whereas secondary RNN has two LSTM layers. The first layer is composed of sequence-to-sequence structure while the second one has sequence-to-label manner.

### C. PARAMETER OPTIMIZATION OF WEIGHTS AND BIASES USING CSO ALGORITHM

CSO is applied in this research to optimize the weights and bias parameters of CLSTM model. Globally, crows are considered as intelligent species in comparison with other birds. It has high potentials and has large-sized brain compared to body. According to brain-to-body theory, brain is marginally lesser in humans. The intelligence of crow is established by massive number of samples. Based on a survey, it has been established that crows have self-experience in mirror tests and skillful in making tools. Crows are capable of remembering faces and it can send warning signals to other crows in case of danger. Also, it makes use of developed tools; share the details, and memorize the secret place of food. Moreover, it observes other birds and chase them to find the secret place of food and grab it, once the bird leaves the nest. Afterwards, crow finds a safe place to store the robbed food so that the actual bird does not find the food. Figure 3 demonstrates the flow chart of CSO. Basically, it uses the knowledge of a thief to speculate a thief's action and selects a secure way to defend its food [20].

Few standards of crows are given herewith.

- It resides in group
- It is capable of remembering the location of food stored in secret places.
- It follows one by one to grab the food.
- It protects their food from being robbed.

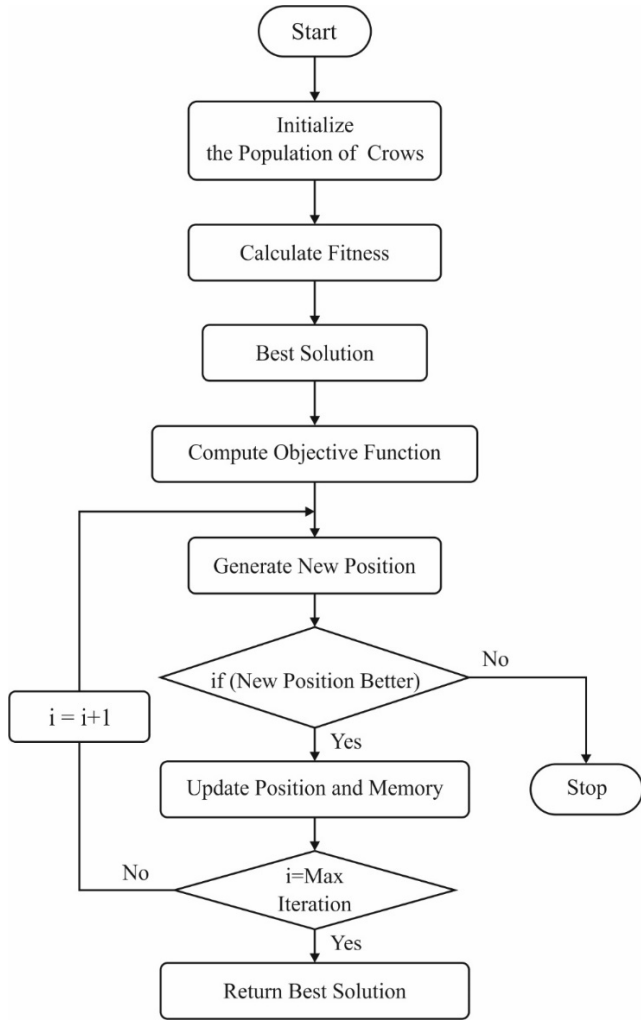


FIGURE 3. Flowchart of CSO algorithm.

Followed by, there are N-dimensional platforms which are composed of massive crows, where  $C$  denotes the overall count of crows and  $u$  defines the position of a crow at time in a Search Space (SS). This has been evaluated in the function given below.

$$V^{u,iter} (p = 1, 2, \dots, C; iter = 1, 2, \dots, iter_{max}) \quad (11)$$

where  $V^{u,iter} = [V_1^{u,iter}, V_2^{u,iter}, \dots, V_c^{u,iter}]$  and  $iter_{max}$  resemble the iterations with higher count. A crow is applicable to remember the place of secret location. At this point, the location of secret place of crow  $u$  is implied as  $s^{u,iter}$ . It is a better location which the crow  $u$  has accomplished. Assume the iteration in which the crow  $v$  requires to be placed in secret location,  $s^{v,iter}$ . At this point, crow  $u$  plans to chase crow  $v$  to the secret place. Here, two processes are carried out as given herewith.

Event 1: Crow  $v$  has no suggestions about which crow  $u$  is chasing. Thus, in the outcome, crow  $u$  reaches the secret position of crow,  $v$ . Then, the new location of crow  $u$  is

developed as given herewith.

$$V^{u,iter+1} = V^{u,iter} + k_j \times fl^{u,iter} \times (S^{v,iter} - V^{u,iter}) \quad (12)$$

where,  $k_j$  implies a random value with uniform distribution between 0 and 1, and  $fl^{u,iter}$  refers to the flight length of crow,  $u$ . The lower value of  $fl$  tends to have local search whereas higher values results in global search.

Event 2: Crow  $v$  understands that crow  $u$  is tracking it. Finally, the theft is prevented and crow  $v$  deceives crow  $u$  by changing its actual position to alternate position of SS. Therefore, events 1 and 2 are illustrated as follows

$$V^{u,iter+1} = \begin{cases} V^{u,iter} + k_j \times fl^{u,iter} \times (S^{v,iter} - V^{u,iter}) k_j \\ \geq AWP^{v,iter} \\ a \text{ random location other wise} \end{cases} \quad (13)$$

where  $AWP^{v,iter}$  implies the awareness of crow  $v$  at iteration.

#### IV. EXPERIMENTAL VALIDATION

This section validates the performance of the presented CSO-CLSTM model in terms of sensitivity, specificity, and accuracy. Besides, the results are examined on heart disease [21] and diabetes [22] datasets under varying number of instances. The presented model was implemented in a PC with specifications such as Motherboard - MSI Z370 A-Pro, Processor - i5-8600k, Graphics Card - GeForce 1050Ti 4GB, RAM - 16GB and File Storage - 1TB HDD.

##### A. RESULTS ON HEART DISEASE DIAGNOSIS

Table 1 shows the classification outcome of CSO-CLSTM model and compares it with existing classifiers with respect to distinct measures on heart disease dataset [21]. When analyzing the results in terms of sensitivity, it is evident that the SVM model demonstrated poor performance over other existing methods. Additionally, the NB-A model attempted to showcase somewhat better sensitivity over SVM. Concurrently, KNN and J48 models resulted in moderately closer and competitive sensitivity values. But the presented CSO-CLSTM model exhibited superior classification performance by obtaining a higher sensitivity value. For example, under 2000 instances, the CSO-CLSTM model reached a maximum sensitivity of 94.80% whereas other models such as KNN, NB-A, SVM, and J48 models accomplished low sensitivity values of 92.60%, 87.90%, 83.20%, and 93.30% respectively. Similarly, under 10000 instances, the proposed CSO-CLSTM method attained a higher sensitivity of 98% while KNN, NB-A, SVM, and J48 methodologies accomplished only least sensitivity values of 93.60%, 89.10%, 84.20%, and 96% respectively.

The results of specificity analysis infers that the SVM approach yielded inferior performance over traditional models. In addition, the NB-A scheme managed to exhibit considerable specificity over SVM. Simultaneously, KNN and J48 frameworks attained acceptable and competing specificity values. However, the newly developed CSO-CLSTM

**TABLE 1.** The performance analysis of existing and proposed method on heart disease dataset.

Sensitivity (%)					
Number of Instances	KNN	NB-A	SVM	J48	CSO-CLSTM
2000	92.60	87.90	83.20	93.30	94.80
4000	88.40	84.60	82.40	92.30	95.20
6000	93.20	86.40	83.90	93.60	96.30
8000	92.40	88.60	82.40	96.90	97.60
10000	93.60	89.10	84.20	96.00	98.00
Specificity (%)					
Number of Instances	KNN	NB-A	SVM	J48	CSO-CLSTM
2000	84.20	83.40	80.20	92.60	94.70
4000	86.10	83.60	82.10	91.20	96.80
6000	87.30	86.90	83.40	92.40	95.00
8000	88.30	82.10	78.40	88.60	91.20
10000	89.30	86.40	84.30	90.40	93.80
Accuracy (%)					
Number of Instances	KNN	NB-A	SVM	J48	CSO-CLSTM
2000	89.40	76.80	73.40	91.60	95.10
4000	91.30	78.60	77.70	92.40	95.90
6000	87.60	77.80	75.60	90.40	95.30
8000	86.40	80.10	78.40	93.20	97.10
10000	89.30	82.40	81.60	92.80	97.40

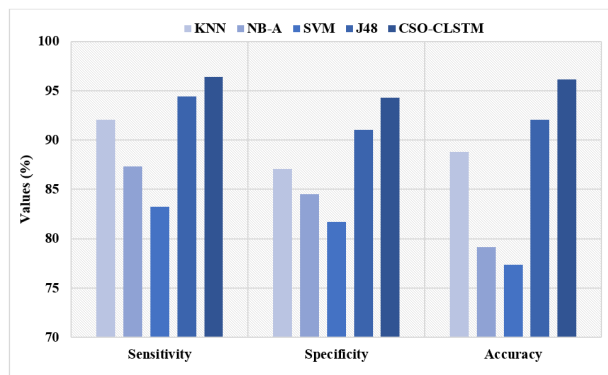
technology accomplished high classification performance and gained optimal specificity value. For instance, under 2000 instances, the CSO-CLSTM scheme achieved a high specificity of 94.70%, whereas the KNN, NB-A, SVM, and J48 technologies accomplished the least specificity values of 84.20%, 83.40%, 80.20%, and 92.60% respectively. In line with these, under the application of 10000 instances, the CSO-CLSTM model attained a supreme specificity of 93.80%, while the KNN, NB-A, SVM, and J48 approaches achieved low specificity values of 89.30%, 86.40%, 90.40%, and 93.80% respectively.

The accuracy analysis inferred that the SVM framework yielded ineffective performance over other traditional approaches. Further, the NB-A scheme managed to produce moderate accuracy over SVM. At the same time, both KNN and J48 approaches produced better and modest accuracy. Thus, the projected CSO-CLSTM approach achieved heavy classification by obtaining the maximum accuracy value. For example, under 2000 instances, the CSO-CLSTM model attained the maximum accuracy of 95.10%, while KNN, NB-A, SVM, and J48 models achieved limited accuracies of 89.40%, 76.80%, 73.40%, and 91.60% correspondingly. At the same time, under the presence of 10000 instances, the CSO-CLSTM technique gained a high accuracy of 97.40%, whereas the KNN, NB-A, SVM, and J48 models obtained only minimum accuracy values of 89.30%, 82.40%, 81.60%, and 92.80% respectively.

Table 2 and figure 4 depict the results of average classification analysis of CSO-CLSTM model on the applied heart disease dataset. The figure clearly portrays that the CSO-CLSTM model outperformed other compared methods by achieving an enhanced average sensitivity of 96.38%, specificity of 94.30%, and accuracy of 96.16%.

**TABLE 2.** The average performance analysis of existing and the proposed CSO-CLSTM method on heart disease dataset.

Measures	KNN	NB-A	SVM	J48	CSO-CLSTM
Sensitivity	92.04	87.32	83.22	94.42	96.38
Specificity	87.04	84.48	81.68	91.04	94.30
Accuracy	88.80	79.14	77.34	92.08	96.16



**FIGURE 4.** Results of average classifier analysis on heart disease dataset.

**B. RESULTS ON DIABETES DATASET**

Table 3 shows the classification result of CSO-CLSTM model with existing classifiers with respect to distinct measures on diabetes disease dataset [22]. When analyzing the results in terms of sensitivity, it is clear that the SVM model demonstrated inefficient performance over other classical methods. Further, both NB-A and KNN model attempted to produce considerable sensitivity over SVM. Meantime, both J48 and FNCA approaches resulted in similar and competitive sensitivity. But the presented CSO-CLSTM model exhibited higher classification performance and obtained high sensitivity value. For example, under 2000 instances, the CSO-CLSTM model achieved a supreme sensitivity of 98.10%, whereas KNN, NB-A, SVM, J48, and FNCA models attained the minimum sensitivity values of 92%, 87.50%, 83%, 93%, and 94.50% respectively. Likewise, under 10000 instances, the CSO-CLSTM model accomplished a maximum sensitivity value of 99.20%, while KNN, NB-A, SVM, J48, and FNCA models accomplished only minimum sensitivity values of 94.20%, 90%, 83.40%, 96%, and 97% respectively.

When analyzing the results with respect to specificity, it is apparent that the SVM model demonstrated poor performance over other existing methods. Besides, both NB-A and KNN models attempted to accomplish improved specificity over SVM. Further, J48 and FNCA models resulted in close and competing specificity. However, the presented CSO-CLSTM model achieved high classification performance and obtained high specificity. For example, under the presence of 2000 instances, the CSO-CLSTM model achieved a high specificity of 98.80%, while the other models such as KNN, NB-A, SVM, J48, and FNCA methodologies attained the minimum specificity values of 84%,

**TABLE 3.** The performance analysis of existing and the proposed method on diabetes disease dataset.

Sensitivity (%)						
Number of Instances	KNN	NB-A	SVM	J48	FNCA	CSO-CLSTM
2000	92.00	87.50	83.00	93.00	94.50	98.10
4000	88.00	86.00	82.50	92.00	93.50	97.50
6000	92.80	88.00	83.80	93.00	94.50	98.90
8000	93.50	88.00	83.00	97.00	98.00	99.40
10000	94.20	90.00	83.40	96.00	97.00	99.20
Specificity (%)						
Number of Instances	KNN	NB-A	SVM	J48	FNCA	CSO-CLSTM
2000	84.00	83.00	82.00	92.50	94.00	98.80
4000	90.00	83.00	83.00	91.00	94.20	97.50
6000	87.00	86.00	83.00	93.00	94.10	96.90
8000	87.50	85.00	80.00	88.00	90.00	94.20
10000	90.00	87.00	84.00	90.50	92.00	97.30
Accuracy (%)						
Number of Instances	KNN	NB-A	SVM	J48	FNCA	CSO-CLSTM
2000	89.00	77.00	74.00	92.00	93.00	95.70
4000	91.00	81.00	76.00	94.00	94.00	97.80
6000	87.00	76.00	75.00	90.00	91.00	96.10
8000	88.00	82.00	78.00	93.50	94.50	98.90
10000	90.00	83.00	80.00	92.50	94.00	97.80

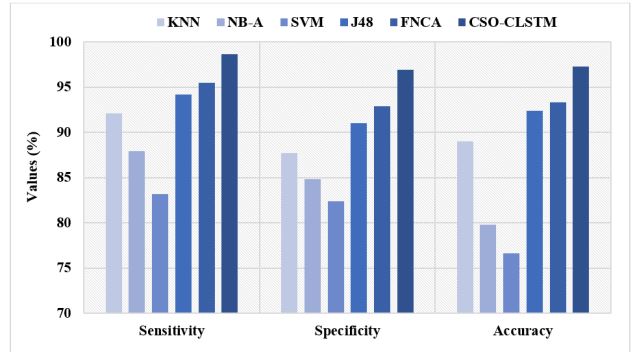
83%, 82%, 92.50%, and 94% respectively. In line with this, under 10000 instances, the CSO-CLSTM model reached high specificity of 97.30%, whereas KNN, NB-A, SVM, J48, and FNCA models achieved only limited specificity values of 90%, 87%, 84%, 90.50%, and 92% respectively.

In accuracy analysis, SVM model demonstrated a low performance over other methods. Additionally, NB-A and KNN models achieved moderate accuracy over SVM. Concurrently, J48 and FNCA models resulted in similar and modest accuracy. But the presented CSO-CLSTM model exhibited superior classification performance and obtained a high accuracy. For example, under 2000 instances, the CSO-CLSTM model accomplished the maximum accuracy of 95.70%, while KNN, NB-A, SVM, J48, and FNCA models produced the least accuracy values of 89%, 77%, 74%, 92%, and 93% respectively. Likewise, under the existence of 10000 instances, the CSO-CLSTM model reached a high accuracy of 97.8%, whereas KNN, NB-A, SVM, J48, and FNCA models accomplished only minimum accuracy values such as 90%, 83%, 80%, 92.50%, and 94% respectively.

Table 4 and figure 5 demonstrate the average classification outcome of CSO-CLSTM method on applied diabetes disease dataset. The figure clearly implies that the CSO-CLSTM approach surpassed the compared models and accomplished an average sensitivity of 98.62%, specificity of 96.94%, and accuracy of 97.26%. From the above-mentioned tables and

**TABLE 4.** The average performance analysis of existing and the proposed CSO-CLSTM method on diabetes disease dataset.

Measures	KNN	NB-A	SVM	J48	FNCA	CSO-CLSTM
Sensitivity	92.10	87.90	83.14	94.20	95.50	98.62
Specificity	87.70	84.80	82.40	91.00	92.86	96.94
Accuracy	89.00	79.80	76.60	92.40	93.30	97.26



**FIGURE 5.** The average classifier results analysis on Diabetes dataset.

figures, it is apparent that the CSO-CLSTM model is effective in performance since during the experimentation, it achieved the maximum accuracy values of 96.16% and 97.26% on heart disease and diabetes diagnoses respectively.

## V. CONCLUSION

The current research work has developed an efficient AI and IoT convergence-based disease diagnosis model for smart healthcare system. The presented model encompasses different stages namely data acquisition, preprocessing, classification, and parameter tuning. IoT devices such as wearables and sensors collect the data while AI techniques utilize the data to perform disease diagnosis. Then, iForest technique is executed to get rid of outliers that exist in the patient data. Followed by, the CSO-CLSTM model is employed to classify the data whether the disease exists or not. In addition, CSO is applied to optimize the weights and bias parameters of the CLSTM model. The utilization of CSO assists in the improvement of diagnostic outcome of CLSTM model. The performance of CSO-LSTM model was validated using healthcare data. During the experimentation, the CSO-LSTM model accomplished a maximum accuracy of 96.16% and 97.26% on heart disease and diabetes diagnoses respectively. This establishes the effectiveness of the presented model. As a part of future scope, the performance can be improved using feature selection techniques which reduce the curse of dimensionality and computational complexity. In addition, the limitations of CSO algorithm such as slow search precision and high possibility of getting into local optima can be resolved with the help of hybrid metaheuristic algorithms.

## REFERENCES

- [1] A. A. Abdellatif, A. Mohamed, C. F. Chiasserini, M. Tili, and A. Erbad, "Edge computing for smart health: Context-aware approaches, opportunities, and challenges," *IEEE Netw.*, vol. 33, no. 3, pp. 196–203, May 2019.



- [2] F. F. Gong, X. Z. Sun, J. Lin, and X. D. Gu, "Primary exploration in establishment of China's intelligent medical treatment," (in Chinese), *Mod. Hos Manag.*, vol. 11, no. 2, pp. 28–29, 2013.
- [3] A. A. Mutlag, M. K. Abd Ghani, N. Arunkumar, M. A. Mohammed, and O. Mohd, "Enabling technologies for fog computing in healthcare IoT systems," *Future Gener. Comput. Syst.*, vol. 90, pp. 62–78, Jan. 2019.
- [4] K. H. Abdulkareem, M. A. Mohammed, A. Salim, M. Arif, O. Geman, D. Gupta, and A. Khanna, "Realizing an effective COVID-19 diagnosis system based on machine learning and IOT in smart hospital environment," *IEEE Internet Things J.*, early access, Jan. 11, 2021, doi: [10.1109/JIOT.2021.3050775](https://doi.org/10.1109/JIOT.2021.3050775).
- [5] A. A. Mutlag, M. K. A. Ghani, M. A. Mohammed, M. S. Maashi, O. Mohd, S. A. Mostafa, K. H. Abdulkareem, G. Marques, and I. De la Torre Díez, "MAFC: Multi-agent fog computing model for healthcare critical tasks management," *Sensors*, vol. 20, no. 7, p. 1853, Mar. 2020.
- [6] G. Villarrubia, J. Bajo, J. De Paz, and J. Corchado, "Monitoring and detection platform to prevent anomalous situations in home care," *Sensors*, vol. 14, no. 6, pp. 9900–9921, Jun. 2014.
- [7] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2399–2406, Aug. 2018.
- [8] A. Kaur and A. Jasuja, "Health monitoring based on IoT using raspberry PI," in *Proc. Int. Conf. Comput., Commun. Autom. (ICCCA)*, Greater Noida, India, May 2017, pp. 1335–1340.
- [9] O. S. Alwan and K. Prahald Rao, "Dedicated real-time monitoring system for health care using ZigBee," *Healthcare Technol. Lett.*, vol. 4, no. 4, pp. 142–144, Aug. 2017.
- [10] U. Satija, B. Ramkumar, and M. Sabarimalai Manikandan, "Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 815–823, Jun. 2017.
- [11] M. Pham, Y. Mengistu, H. Do, and W. Sheng, "Delivering home healthcare through a cloud-based smart home environment (CoSHE)," *Future Gener. Comput. Syst.*, vol. 81, pp. 129–140, Apr. 2018.
- [12] M. S. Hossain and G. Muhammad, "Deep learning based pathology detection for smart connected healthcare," *IEEE Netw.*, vol. 34, no. 6, pp. 120–125, Nov. 2020.
- [13] M. Z. Uddin, "A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system," *J. Parallel Distrib. Comput.*, vol. 123, pp. 46–53, Jan. 2019.
- [14] S. S. Ram, B. Apduhan, and N. Shiratori, "A machine learning framework for edge computing to improve prediction accuracy in mobile health monitoring," in *Proc. Int. Conf. Comput. Sci. Appl.* Cham, Switzerland: Springer, Jul. 2019, pp. 417–431.
- [15] J. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, "Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks," in *Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2019, pp. 601–604.
- [16] I. Azimi, A. Anzanpour, A. M. Rahmani, T. Pahikkala, M. Levorato, P. Liljeberg, and N. Dutt, "HiCH: Hierarchical fog-assisted computing architecture for healthcare IoT," *ACM Trans. Embedded Comput. Syst.*, vol. 16, no. 5s, p. 174, Oct. 2017.
- [17] G. Muhammad, M. S. Hossain, and N. Kumar, "EEG-based pathology detection for home health monitoring," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 603–610, Feb. 2021, doi: [10.1109/JSAC.2020.3020654](https://doi.org/10.1109/JSAC.2020.3020654).
- [18] J. Ren, J. Guo, W. Qian, H. Yuan, X. Hao, and H. Jingjing, "Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms," *Secur. Commun. Netw.*, vol. 2019, pp. 1–11, Jun. 2019.
- [19] N. Michielli, U. R. Acharya, and F. Molinari, "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals," *Comput. Biol. Med.*, vol. 106, pp. 71–81, Mar. 2019.
- [20] S. Manimurugan, A.-Q. Majdi, M. Mohammed, C. Narmatha, and R. Varatharajan, "Intrusion detection in networks using crow search optimization algorithm with adaptive neuro-fuzzy inference system," *Microprocessors Microsyst.*, vol. 79, Nov. 2020, Art. no. 103261.
- [21] *Heart Disease Data Set*. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/heart+disease>
- [22] *Pima Indians Diabetes Database*. [Online]. Available: <https://www.kaggle.com/uciml/pima-indians-diabetes-database>



**ROMANY FOUAD MANSOUR** received the B.Sc. and M.Sc. degrees in computer science from Assiut University, Egypt, in 1998 and 2006, respectively, and the Ph.D. degree from the University of Assiut, in 2009. He is currently working as an Associate Professor with the Department of Mathematics, Faculty of Science, New Valley University, Egypt. His research interests include pattern recognition, computer vision, computer networks, soft computing, image processing, evolutionary computation, and machine learning.



**ADNEN EL AMRAOUI** received the Ph.D. degree in process control from the Technical University of Belfort-Montbéliard (UTBM), France, in 2011. From September 2013 to August 2014, he was a Postdoctoral Researcher with the French Engineering School "Ecole Centrale de Lille (EC-Lille)." From September 2014 to August 2018, he was an Associate Professor with the University of Orléans, France. Since September 2018, he has been an Associate Professor with the University of Artois, France. He is the first and the corresponding author of several publications in refereed journals, such as *COR*, *CIE*, *EEAI*, *EJIE*, *4-OR*, *JSSSE*, *SIC*, and so on, and in several IEEE and IFAC conferences. His research interests include combinatorial optimization, artificial intelligence, approximation, interval analysis, scheduling and production planning, supply chain management, transportation systems, and maintenance.



**ISSAM NOUAOURI** received the Ph.D. degree in industrial and logistic engineering from the University of Artois. He was an Engineer in industrial engineering with the Ecole National d'Ingénieurs de Tunis. He is currently an Associate Professor with the University of Artois and the LGI2A Laboratory. He has also been administrative and scientific Manager in research projects (UTIQUE "France-Tunisia," BQR-University of Artois). He is also developing research focusing on logistics and operational research in health care. He was on the 2012 lists of University Lecturers in Automatic and Computer Science (CNU61). He is a Manager of the Bachelor's degree in Business and Distribution with the University of Artois. He is the author of several international publications in academic journals, such as *ORHC*, *JC*, *JAOR*, and so on, and in international conferences, such as IEEE, IFAC, and so on. He is a member of scientific committee to review paper in international conference, such as GISEH, MOSIM, and so on, and journals, such as *International Journal of Operations and Production Management*, *International Journal of Manufacturing Technology and Management*, and so on.



**VICENTE GARCÍA DÍAZ** graduated in prevention of occupational risks. He received the Ph.D. degree in computer engineering from the University of Oviedo. He is currently an Associate Professor with the Computer Science Department, University of Oviedo. He is a Certified Associate in Project Management through the Project Management Institute. His research interests include model-driven engineering, domain specific languages, technology for learning and entertainment, project risk management, and software development processes and practices.



**DEEPAK GUPTA** (Senior Member, IEEE) received the Ph.D. degree from Dr. APJ Abdul Kalam Technical University. He held a post-doctoral position with Intel, Brazil. He is currently working as an Assistant Professor with the Maharaja Agrasen Institute of Technology (GGSIPU), Delhi, India. He is an eminent academician and plays versatile roles and responsibilities juggling between lectures, research, publications, consultancy, community service, Ph.D.

and Postdoctoral Supervision, and so on. With 13 years of rich expertise in teaching and two years in industry; he focuses on rational and practical learning. He has contributed massive literature in the fields of human-computer interaction, intelligent data analysis, nature-inspired computing, machine learning, and soft computing. He has actively been part of various reputed International conferences. He is not only backed with a strong profile but his innovative ideas, research's end-results and notion of implementation of technology in the medical field is by and large contributing to the society significantly. He has authored/edited 46 books with national/international level publisher, such as Elsevier, Springer, Wiley, and Katson. He has published 157 scientific research publications in reputed international journals and conferences, including 78 SCI indexed journals of the IEEE, Elsevier, Springer, Wiley, and many more. He has served as the Editor-in-Chief, a Guest Editor, an Associate Editor in SCI and various other reputed journals, such as Elsevier, Springer, Wiley, and MDPI. He is the Editor-in-Chief of *OA Journal-Computers* and *Quantum Computing and Applications (QCAA)*. He is an Associate Editor of *Expert Systems (Wiley)*, *Intelligent Decision Technologies (IOS Press)*, the *Journal of Computational and Theoretical Nanoscience*, an Honorary Editor of the *ICSES Transactions on Image Processing and Pattern Recognition*. He is also a Series Editor of Elsevier Biomedical Engineering (Elsevier), Intelligent Biomedical Data Analysis (De Gruyter, Germany), and Explainable AI (XAI) for Engineering Applications (CRC Press). He is appointed as a Consulting Editor at Elsevier. He was invited as a faculty resource person/session chair/reviewer/TPC member in different FDP, conferences, and journals.



**SACHIN KUMAR** (Senior Member, IEEE) received the Ph.D. degree in data mining/machine learning from the Indian Institute of Technology Roorkee, India, in 2017. He worked as Postdoctoral Fellow and a Senior Researcher with the National University of Science and Technology, MISIS, Moscow, Russia. He is currently an Associate Professor and a Senior Research Fellow with the Department of Computer Science, South Ural State University, Chelyabinsk, Russia. His

main research interests include machine learning applications in materials discovery, healthcare, the IoT, medicine, transportation, and NLP. He is also serving as a reviewer and a guest editor for several reputed journals.

• • •