**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 30, No. 2, May 2023, pp. 846~859
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v30.i2.pp846-859


# A remote health monitoring framework for heart disease and diabetes prediction using advanced artificial intelligence model

**Adari Ramesh, Ceeke Kalappagowda Subbaraya, Ravi Kumar Guralamata Krishnegowda**
Department of Computer Science and Engineering, College of BGS Institute of Technology (BGSIT),
Adichunchanagiri University (ACU), Karnataka, India

**Article Info**

***Article history:***

Received Jun 23, 2022
Revised Dec 26, 2022
Accepted Dec 28, 2022

***Keywords:***

Artificial intelligence
Hungarian and PIMA datasets
Preprocessing
Probabilistic guidednaïve distribution classification
Remote health monitoring
Smart devices
Unified levy modeled crow search optimization

**ABSTRACT**

Remote health monitoring frameworks gained significant attention due to their real intervention and treatment standards. The proposed work intends to design an artificial intelligence (AI) based remote health monitoring framework for predicting heart disease and diabetes from the given medical datasets. In this framework, the smart devices are used to gather the health information of patients, and the obtained information is integrated together by using different nodes that includes the detecting node, visualization node, and prognostic node. Then, at that point, the health care dataset preprocessing is performed to standardize the characteristics by recognizing the missing qualities and taking out the unessential characteristics. Consequently, the unified levy modeled crow search optimization (ULMCSO) algorithm is employed to select the optimal features based on the global fitness function, which helps increase the accuracy and reduce the training time of the classifier. Finally, the probabilistic guided naïve distribution (PGND) based classification model is utilized for predicting the label as to whether normal or disease affected. During an evaluation, two different datasets, such as PIMA and Hungarian, are used to validate and compare the results of the proposed model by using various performance measures.



***Corresponding Author:***

Adari Ramesh
Department of Computer Science and Engineering, College of BGS Institute of Technology (BGSIT)
Adichunchanagiri University (ACU)
B. G. Nagara, Nagamangala, Karnataka, India
Email: adariramesh3@gmail.com

## 1. INTRODUCTION

Public health monitoring is one of the most essential and fundamental concerns that need to be focused on to prevent people from health hazards [1]. In current days, most people are highly affected by heart disease and diabetes, which are considered life-threatening diseases because it affects the other parts of the nerve system and kidney. Due to the low-quality treatments and inefficient clinical diagnosis, most healthcare systems [2], [3] could use the decision making systems for proper disease prediction and diagnosis. Because earlier disease identification helps to prevent people from harmful diseases. Recently, developing an automated remote monitoring system is more vital for observing the healthcare status [4], [5] of patients, which is accomplished by using health devices, smartwatches, and smart mobile phones. The primary purpose of using the remote monitoring system is to provide proper treatments to the patients during emergency cases. For this purpose, an intelligent machine learning-based smart health framework is constructed in this work, which helps to consolidate the patient data [6] from the smart devices for early detection of risks and keep the healthcare professionals updated with the present health information patients. In this framework, the detecting node, visualization node, and prognostic node have been used to gather the health information from patients for generating the alert to the healthcare professionals [7]. This pipeline can be implemented with the modules of data imputation, optimization and classification. Here, the data imputation is performed for scaling the data obtained from the medical datasets, and it helps to improve the quality of patterns. After that, the feature selection can be done by using the novel optimization technique to reduce the dimensionality of data to improve the prediction rate. Finally, an advanced artificial intelligence (AI) based machine learning technique can be implemented for accurately predicting the given data into normal or abnormal classes. The main contributions of this paper are as follows:

– To obtain the health information of patients from the smart devices, three different nodes such as detecting node, visualization node, and prognostic node have been utilized in this framework.
– To fine-tune the given input datasets by eliminating the noise and normalizing the attributes, the dataset preprocessing is performed initially.
– A unified levy modeled crow search optimization (ULMCSO) based feature selection algorithm is employed to select the features based on the global fitness value with reduced iterations.
– The probabilistic guided naïve distribution (PGND) classification algorithm is deployed to predict the classified output as normal accurately or disease affected.
– To validate the efficiency and results of these approaches, the two different and most popular datasets, such as Hungarian and PIMA, are utilized in current system.

The overflow bits of this paper are structuralized into the going with pieces: Section 2 takes apart the continuous social occasion, movement, and solicitation perspectives utilized for somewhat investigating the success status of patients. Section 3 presents the working systems for the proposed work with its general stream and estimations. Section 4 assesses the presentation and near examination of existing and proposed procedures utilizing different measures. At last, the general paper is summed up with its future work in section 5.

## 2. RELATED WORKS

This segment surveys a portion of the regular works connected with information bunching, enhancement, AI, and profound learning methods used to foster a far off medical care checking framework for nursing patients with various kinds of infections. Moreover, it analyzes the benefits and obstacles of the ongoing procedures as shown by their basic features and used norms. Gondalia *et al.* [8] introduced an automated internet of things (IoT) based health monitoring system by using the machine learning technique for observing the health status of war soldiers. The main intention of this paper was to accurately identify the location of soldiers for monitoring their health status, who injured in the battlefield. Here, the health of soldiers were monitored by using the GPS control, and sensors like (heartbeat and temperature). Moreover, the k-means clustering algorithm was utilized in this work for analyzing the information obtained from the sensors. According to the sensor inputs, the different types of actions like sitting, walking or running in case of wound or blasting have been accurately predicted by clustering the attributes. However, the k-means clustering was not more suitable for this kind of real time applications due to its reduced efficiency, global clustering, and varying initial partitions. Malasinghe *et al.* [9] presented a comprehensive survey for the selecting the suitable technique to develop a remote patient health monitoring system. Also, the different types of security issues associated to the e-health systems were discussed with its appropriate solutions. Ramkumar *et al.* [10] designed a new remote patient monitoring framework using animproved machine learning technique. The pilot data has been utilized in this system for analysis, where the wearable technology could be used to monitor the health status of patients. Chatrati *et al.* [11] encouraged a canny home prosperity really taking a look at framework for affirmation the diabetes and circulatory strain patients from a good way.

The principal reason for this work was to dissect the glucose readings of the patients at their home for giving the medical care office on the off chance that anomaly. Moreover, the circulatory strain was likewise identified by utilizing the mix of direction and AI models. Here, the support vector machine (SVM) request strategy was utilized to arrange the diabetic and non-diabetic patients considering the data planning set. Moreover, the model optimization algorithm was utilized to optimally select the features for improving the accuracy of classification, which also helps to reduce the overfitting. Still, this work requires to develop the proper graphical user interface for automatically transferring the medical status of patients to the doctors at the time. Vitabile *et al.* [12] utilized a smart health monitoring system for analyzing the psychological conditions and health status of patients based on the measures of chest sounds, temperature, blood pressure, heart rate, and electrocardiograms (ECG). Here, the blockchain methodology has been utilized to ensure the privacy and security of medical data. The primary advantage of this system was, it has better ability to handle the large dimensional data with reduced time delay. Nair *et al.* [13] employed a spark-based machine learning methodology for remotely predicting the health status of patients. Here, the decision tree classification mechanism was utilized to predict the disease based on the attributes obtained from the dataset. Also, it was more capable for handling the huge datasets based on the partitioning of feature models. Moreover, this classification model could efficiently reduce the generalization error and overfitting value. Still, it follows some computational steps for predicting the classified labels, which was the key limitation of this work.

Li *et al.* [14] implemented a multi-stationary approach for remotely monitoring the health status of older people. In this framework, two different types of sensor have been utilized to predict the disease, which includes radar and wearable sensors. For improving on the cycle and expanding the exactness of grouping, different element choice models have been used in this work, which incorporates the channel strategies, covering techniques, and implanted strategies. Besides, the K-nearest neighbor (KNN) and SVM request approaches have been used for predicting an infection, and the got consequences of these methods were looked at in light of the proportions of precision and handling time. According to the validation, it was analyzed that the SVM technique outperforms the KNN model with improved performance results. Kaur *et al.* [15] deployed a random forest classification technique for designing an efficient health monitoring system to remotely monitor the health status of patients. Here, the IoT based healthcare system was developed for predicting chronic diseases at the time of medical emergencies. Based on the types of various disease, the performance of classification techniques such as KNN, linear SVM, random forest, decision tree, and multilayer perceptron were verified and compared by using the measures of accuracy, area under curve (AUC), accuracy, and review. In view of the acquired outcomes [16], it was analyzed that the random forest technique outperforms the other techniques with increased detection accuracy.

Verma and Sood [17] presented a haze helped IoT empowered structure for remotely checking the medical care information through the brilliant passages. Here, the event triggering data transmission methodology could be used for analyzing the patient's health data based on the temporal health index. Azimi *et al.* [18] implemented an edge based deep learning methodology for designing a hierarchical healthcare monitoring framework. The primary motivation behind this work is to guarantee the expanded accessibility and exactness of order at the hour of ongoing checking. In this system, the medical care give could acquire the clinical information from the cloud server through the sensor gadgets associated in the organization. The primary advantage of this work was, it efficiently reduced the response time by properly monitoring the health information. Moshin *et al.* [19] presented a multilayer taxonomy for developing an efficient real time remote health monitoring system. Here, the wireless body area network (WBAN) telemedicine architecture model was developed for monitoring the patient's health information using various body sensor devices. It includes blood pressure sensor, ear sensor, motion sensor, electroencephalogram (EEG), ECG and electromyogram (EMG) sensors, which gathers the medical information from the human body and transmits it to the healthcare professional through wireless technology.

Devarajan *et al.* [20] constructed an energy efficient fog assisted health monitoring framework for remotely monitoring the diabetic patients. The main purpose of this work was to reduce the latency, complexity, and improve the energy efficiency of healthcare support system. Hassan *et al.* [21] introduced the remote pain monitoring architecture using an IoT incorporated Fog technology for e-healthcare system. This paper objects to accomplish the reduced execution cost, time delay, and resource consumption required for developing an efficient web application. Moreover, the different types of e-healthcare services have been validated in this work, which includes robotic services, audio/video communication, and monitoring application. In this framework, the fog node gathers the patient health information by using the actuators and sensors, and the obtained data were transferred to the cloud storage using the proxy server. This framework works in light of the early bird gets the worm planning process, subsequently it was not more appropriate for all sort of health-related crisis applications, which was the significant restriction of this work. Hasan *et al.* [22] sent a fix based continuous medical care observing framework utilizing an IoT innovation. The principal motivation behind this work was to create an exceptionally solid eHealth system for lessening both the expense and human exertion. Here, the low power radio recurrence innovation could be utilized for laying out the information transmission with diminished piece mistake rate Table 1. Contains the accuracy of the different algorithems used and utilization of the IoT devise.

Here, the SVM grouping method was used to classify Based on the survey, it is examined that the current works are exceptionally centered around fostering a far-off medical services framework for observing the wellbeing status of patients utilizing the savvy gadgets like sensors, shrewd watches and other brilliant gadgets. Still, it faced the following limitations:

- Inefficient data transmission between the sensors and storage components.
- Increased time delay of processing.
- High bit error rate and complexity in classification.
- Reduced system performance.

Hence, the proposed work objects to design an efficient remote health monitoring framework for tracking the health status of patients using an advanced AI mechanism.

Table 1. Comparative analysis between the existing algorithms

| References | Technique | Description | Utilization of IoT device | Accuracy |
|---|---|---|---|---|
| Makhadmeh and Tolba [23] | Deep belief neural network (DBNN) | An IoT weable medical device has been utilized in this model for data gathering. | Yes | 99% |
| Mohan *et al.* [24] | Random forest (RF) integrated with linear model | It objects to predict the cardio vascular disease. | No | 88.9% |
| Nalluri and Roy [25] | Multi-objective optimization-based classification | A hybrid disease detection framework is developed. | No | 94% |
| Haq *et al.* [26] | Intelligent machine learning model | A hybridized framework is developed for predicting the heart disease. | No | 89% |
| Uyar and İlhan [27] | GA integrated with recurrent fuzzy NN | It intends to design a heart disease detection framework. | No | 97.8% |
| Ahmed [28] | K-nearest neighbor (KNN) Classification | It developed an IoT based heart rate monitoring framework. | Yes | 96% |
| Nazari *et al.* [29] | Fuzzy AHP model | A likelihood-based disease detection system is developed. | No | 95% |
| Vivekanandan and Iyengar [30] | Differential evolution (DE) integrated with fuzzy NN | It objects to implement the machine learning classifier for heart disease prediction. | No | 83% |
| Ali *et al.* [31] | Statistical DNN classification | It intends to develop a heart disease detection system. | No | 93.3% |
| Khiarak *et al.* [32] | Meta-heuristic optimization-based classification | The cardiac disease diagnosis system is developed. | No | 94% |

## 3. PROPOSED METHOD

This part examines about the functioning model of proposed far off wellbeing checking framework utilizing progressed enhancement and arrangement approaches. The principal commitment of this paper is to foster a wellbeing checking system for partner the clinical data got from the patients utilizing the individual savvy gadgets. In light of this clinical information, the diabetes dangers and heart chances are anticipated in before for giving clinical updates to the patients. The proposed smart healthcare monitoring framework and its components are shown in Figure 1. Here, the information like patient name, age, gender, height/weight, heart rate, blood pressure, glucose level, calories, medication level, and blood oxygen have been gathered using the sensor nodes. In this framework, the one-time authentication is performed for fusing the medical information obtained from the sensors. Then, the mobile application can directly extract the required data from the cloud platforms, hence the user does not required to enter their data, which is one of the key benefit of this system. For this purpose, the GoogleFit cloud service is utilized in this model that maintains the medical information of patients obtained through smart devices. Moreover, the application programming interfaces (API) has been utilized to establish the connection and data transmission between the end-user devices and software arbitrators.

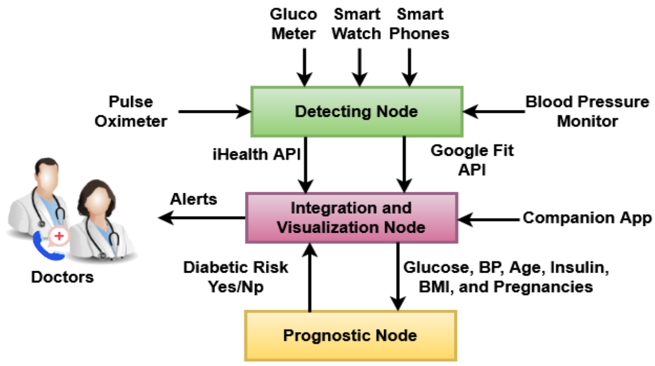



Figure 1. Smart remote healthcare framework

The different types of nodes exist in this framework are as follows:

a) Detecting node – The main functionality of this node to collect the patient medical information from the smart devices.
b) Visualization node – It combines all the medical information obtained from various sources for making it as accessible in the web portal.
c) Prognostic node – In which, the machine learning classifier is utilized to predict the disease according to the subset of features extracted from the patient data.

The working flow of the proposed methodology is shown in Figure 2, and its corresponding architecture is depicted in Figure 3, which holds the following modules:
– Data preprocessing and normalization
– ULMCSO based feature selection
– PGND classification

Initially, the data preprocessing and normalization are performed to improve the quality of dataset by identifying the missing values, and eliminating the irrelevant information. Because, the disease prediction results are highly depending on the attributes of dataset, hence it must be fine-tuned for accurate detection and classification. After that, the feature scaling is performed to standardize the values of attributes in the given dataset. Consequently, the ULMCSO algorithm is applied to optimally select the features for reducing the dimensionality and increasing the accuracy of classifier. Then, the selected features are used to train the classifier for predicting the disease of patients with reduced computational complexity and increased detection efficiency. Here, the PGND based classification technique is implemented to accurately detect the disease based on the features selected from the given dataset.

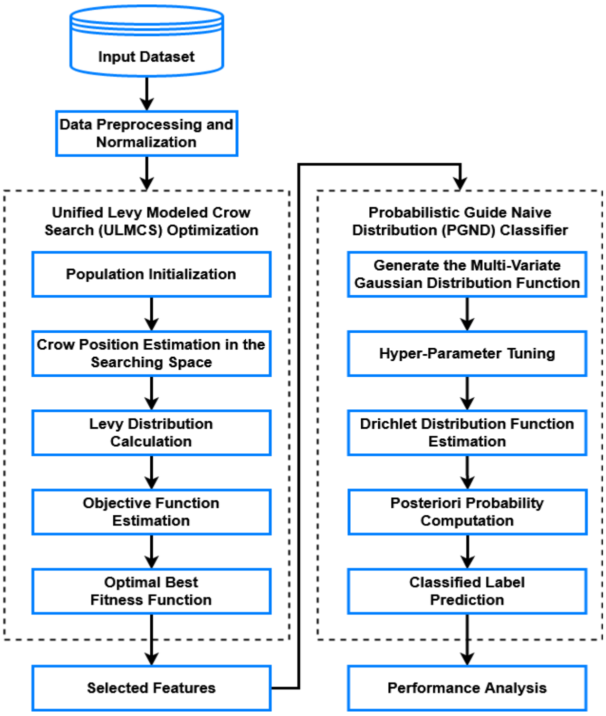



Figure 2. Working flow of the proposed health monitoring

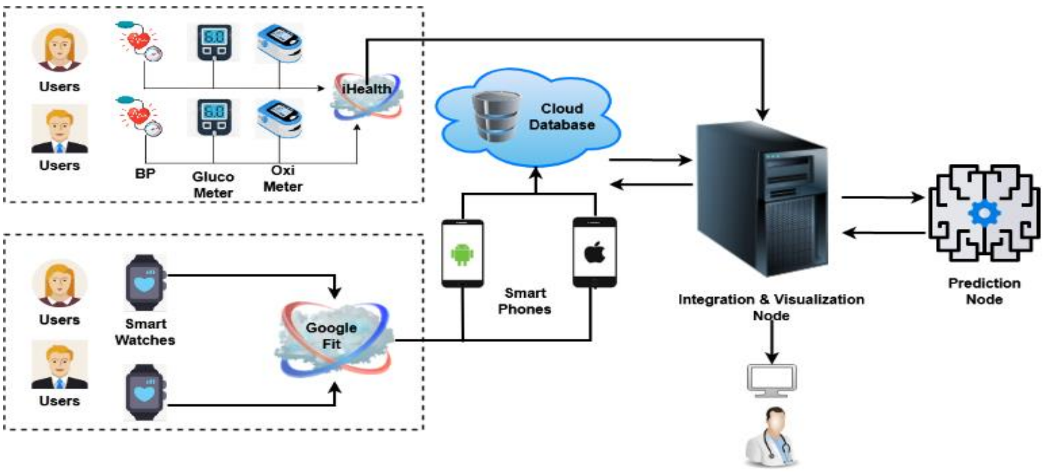



Figure 3. Architecture model of the proposed system

### 3.1. Dataset normalization and preprocessing

In any case, the dataset preprocessing and normalization processes are performed for tuning the data by recognizing the missing characteristics and killing the unessential properties. Besides, the missing information substitution and clamor evacuation are considered as the fundamental activities in dataset preprocessing, in light of the fact that the commotion free information assists with acquiring a superior identification rate. Here, the input dataset is examined initially for analyzing the number of data to identify the missing attributes based on the median value. During this cycle, the information is organized in a rising request for processing the middle worth. Thus, the irrelevant qualities and missing qualities are supplanted with the assessed middle worth. After that, the data normalization is performed by standardizing the values in the range of 0 to 1, which helps to evaluate the patterns of disease. For data normalization, the standard deviation is estimated by using the regression model as shown in (1) and (2).

$$D = \rho_0 + \rho_i S + \varepsilon_i \quad \text{for } i = 1, 2 \dots n \tag{1}$$

$$D_i = \rho_0 + \rho_i S + \varepsilon_i^* \tag{2}$$

By using these mathematical models, the residual value is computed. Where, D indicates the pair of data, ρ_0 and ρ_i are defined as the least square values, S is the input data, and ε_i denotes the error value. The average values are estimated for the sample data by using the standard deviation as shown in (3):

$$\mu = \frac{\sum_{i=1}^{n} S_i}{F_d} \tag{3}$$

where, $\sigma$ indicates the standard deviation, $S_i$ represents the input data and $F_d$ is the frequency of data. Consequently, the data normalization is performed as shown in (4) and (5),

$$D_N = \frac{\varepsilon_i^*}{\dot{\sigma_i}} \tag{4}$$

$$D_N = \frac{S_i - \mu_i^*}{\dot{\sigma_i}} \tag{5}$$

where, $D_N$ indicates the normalized data, $\varepsilon_i^*$ is the residual value, and $\dot{\sigma_i}$ defines the variance. Then, this preprocessed data can be used for the optimization and classification processes.

### 3.2. ULMCSO based feature selection

Close to the completion of preprocessing, the ideal number of components are browsed the dataset by using the proposed ULMCSO computation, which helps with diminishing the dimensionality of features. The primary motivation behind applying the ULMCSO procedure is to prepare the information model of classifier by utilizing the ideal number of highlights. Typically, the feature selection algorithms are mainly in the prediction or classification application systems During this cycle, the information is organized in a rising request for processing the middle worth. Thus, the irrelevant qualities and missing qualities are supplanted with the assessed middle worth. Increasing the overall prediction performance. Here, the optimization is performed by incorporating the functionalities of levy flight and CSO mechanisms. The proposed ULMCSO is a kind of meta-heuristic optimization technique, which is used to identify the global best fitness function for optimally choosing the most reasonable highlights. The principal advantages of utilizing this strategy are as per the following:

– Increased convergence speed
– Reduced future dimensionality
– Minimal time consumption for classifier training and testing
– Increased classification accuracy

Due to the food hiding ability and memory capacity, the crows are generally considered as an intelligent bird that uses its own knowledge for food collection and storage. Also, the CSO streamlining approach is widely used in numerous application frameworks for settling the complex multi-objective enhancement issues. Because of its wasteful looking through capacity and space, the CSO strategy limits with the issue of decreased combination rate. Thus, the toll flight procedure is consolidated with the CSO calculation for further developing the general intermingling velocity, and precision. Moreover, the hybrid ULMCSO can identify the best fitness value with reduced number of iterations. In this mechanism, the set of populations $F_n$ are randomly initialized at first with N number of crows. Here, the maximum count for iteration is considered as $m_k$, and the position of $c^{th}$ crow having the dimensional searching space *d*and iteration *t* is computed by using the (6):

$$Cr_{c,t} = [Cr_{c,t}^1, Cr_{c,t}^2 \dots Cr_{c,k}^n] \quad (6)$$

where, c = 1, 2 … $F_n$ and $t$ = 1, 2 …$m_t$. In this model, each crow has the ability to evoke the location or position for hiding the food source, which accomplished before starting the next iteration as estimated as (7).

$$H_{c,t} = [h_{c,t}^1, h_{c,t}^2 \dots h_{c,t}^n] \quad (7)$$

Moreover, the crow $c$can use the random path for safeguarding the food source from the other crow, which is represented the random path selection of optimization. In this integrated algorithm, the random path is identified by using the levy flight algorithm, where the searching probability is predicted for analyzing the behavior of crow. Let, consider that the random number $W_c$ is uniformly distributed between the range of 0 to 1, which is represented as (8).

$$W_c = Levy \sim x = a^{-\lambda}, \text{ Where } (1 < \lambda \leq 3) \quad (8)$$

This operation is explained by using the follow model (9),

$$Cr_{c,t+1} = \begin{cases} Cr_{c,t} + W_c \cdot Am_{c,t} \cdot (Be_{v,t} - Cr_{v,t}) W_c \geq AP \\ Random\ placement \quad\quad\quad Otherwise \end{cases} \quad (9)$$

where, $Am_{c,t}$ defines the amplitude of crow $c$, $Be_{v,t}$ denotes the identified best possible solution of crow $v$, and $W_c$ is the random number. Finally, the memory vector is updated by using the (10):

$$Cr_{c,t+1} = \begin{cases} Cr_{c,t+1} \quad if\ O(Cr_{c,t+1})\ if\ better\ than\ (Be_{c,t}) \\ Be_{c,t} \quad\quad\quad Otherwise \end{cases} \quad (10)$$

where, $O(.)$ indicates the objective function. By using this value, the best optimal function is identified for selecting the features to train the classifier.

Crow search algorithm 1 is a new type of swarm intelligence optimization algorithm proposed by simulating the crows' intelligent behavior of hiding and retrieving deseased the algorithm has the characteristics of simple structure, few control parameters, and easy implementation

*Unified Levy Modeled Crow Search Optimization (ULMCSO) – Algorithm-I*

*Input:* Preprocessed Dataset;
*Output:* Optimal selection of features;
*Step 1:* Initialize the set of populations $F_n$, Number of crows N, maximum count for iteration $m_k$, dimensional searching space d and iteration t;
*Step 2:* The population initialization with n number of crows with t iteration $Cr_{c,t}$is represented in equ (1);
*Step 3:* Then, the location or position of hiding food source by the crow c is estimated by using the equ(7);
*Step 4:* Generate the random number between the range of [0, 1] by using the levy flight modeling as shown in equ (8);
*Step 5:* Then, its random placement $Cr_{c,t+1}$ is updated by using equ (9);
*Step 6:* Update the memory function according to the amplitude $Am_{c,t}$, best possible solution $Be_{v,t}$, and random number $W_c$.
*Step 7:* Based on the objective function $O(.)$, the best optimal solution is computed for optimally selecting the number of features;

### 3.3. PGND classification

Subsequent to smoothing out, the classifier is ready by using the best game plan of components that helps with chipping away at the accuracy and diminishes the time use of taking care of. In this work, the PGND order system is carried out to recognize the illness from the given datasets by utilizing the ideal number of chosen highlights. It is a kind of probabilistic mixture model mainly used for labeling the dataset based on the Gaussian distribution value. In this algorithm, the model parameters are selected by using the maximum likelihood function obtained from the available data. During this process, the sample mean and covariance values are computed for predicting the class with increased accuracy. The key benefit of using this After streamlining, the classifier is prepared by utilizing the ideal arrangement of elements that assists with working on the precision and decreases the time usage of handling. In this work is it effectively avoids the data over training problem with increased robustness. Also, it is more suitable for handling the large dimensional datasets with reduced time consumption and error rate. Because, the hyper parameter tuning is performed with zero mean, and unit variance measures. Initially, the multi-variate distribution function is computed with $T$ number of Gaussian distributions as shown in (11):

$$p(opt_i|ch_i = t) = N(Me_t, \mathrm{Co}_t) \tag{11}$$

where, $opt_i$ is the optimized dataset, $ch_i$ indicates the current operating condition, $Me_t$ is the mean value, and $\mathrm{Co}_t$ defines the covariance. Also, there are $T$number of parameters used to define the feature space model of $\{(Me_1, \mathrm{Co}_1), (Me_2, \mathrm{Co}_2) \ldots. (Me_t, \mathrm{Co}_t)\}$. Consequently, the statistical model of mixing properties is estimated by using the (12):

$$\Theta = \{(Me_1, \mathrm{Co}_1, \Gamma_1), (Me_2, \mathrm{Co}_2, \Gamma_2) \ldots (Me_t, \mathrm{Co}_t, \Gamma_t)\} \tag{12}$$

where, $\Gamma_t$ indicates the mixing proportions of class $t \in S$, $\Gamma = \{\Gamma_1, \Gamma_2 \ldots \Gamma_t\}$. Then, the normal inverse wischart distribution $NIW$ is computed based on the conjugate of distribution as shown in (13).

$$p(Me_t, Co_t) = NIW(mk_0, r_0, q_0, MV_0) \tag{13}$$

Where, $mk_0$ is the prior mean of $Me_t$, $MV_0$ is the prior mean of $Co_t$, $r_0$ and $q_0$ indicates the strength of prior. Consequently, the identity matrix is constructed [ID × ID] with the *d*dimensional vector, and the distribution over the labelled space is denoted by using the (14).

$$p(\Gamma) = Dir(\omega)\omega \prod_{t=1}^{T} \Gamma_t^{\omega_t - 1} \tag{14}$$

Where, the hyper parameters $\omega = \{\omega_1, \ldots \omega_T\}$ that is used to integrate the posterior probability of each class. Then, its equally weighted factor $\omega_t = \frac{n}{T}, \forall t$ is computed for each class, and the generative statistical model $p(ch_i, opt_i, \Theta)$ is also determined. Consequently, the labelled data $LD_d$ is used to generate the initial number of classes $T$, where the model parameters are estimated by using the Bayesian function. Moreover, the posteriori probability is estimated and updated with the model parameters as shown in (15).

$$p(Me_t, Co_t|ch_i = t, LD_d) = NIW(mk_n, r_n, q_n, MV_n) \tag{15}$$

The parameters $mk_0, r_0, q_0, MV_0$ are computed as following (16)-(19):

$$mk_n = \frac{r_0}{r_0 + n_t} mk_0 + \frac{n_t}{r_0 + r_t} \overline{opt}_t \tag{16}$$

$$r_n = r_0 + n_{ch} \tag{17}$$

$$q_n = q_0 + n_t \tag{18}$$

$$MV_n = MV_0 + MV + r_0 mk_0 mk_0^S - r_n mk_n mk_n^S \tag{19}$$

where, $n_t$ indicates the number of observations in the labelled data $LD_d$, and $\overline{opt}_t$ defines the sample mean with the label *t*. Then, the sum of square matrix computed for each class *t* is shown in (20).

$$MV = \sum_{i=1}^{n_{ch}} opt_i opt_i^S \tag{20}$$

The posterior probability is estimated based on the categorical distribution of the dirichlet function as shown in (21),

$$p(\Gamma|LD_d) \propto \prod_{ch=1}^{T} \Gamma_{ch}^{n_{ch} + \omega_{ch} - 1} \tag{21}$$

according to the posteriori distribution function, the Bayes rule is applied to predict the classified label from the given data as shown in (22).

$$p\left(\widehat{ch}_i = t \middle| \widehat{opt}_i, LD_d\right) = \frac{p(\widehat{opt}_i|\widehat{ch}_i = t, LD_d) p(\widehat{ch}_i = t|LD_d)}{p(\widehat{opt}_i|LD_d)} \tag{22}$$

By utilizing this model, the characterized mark is anticipated as whether typical or illness impacted. The significant advantages of utilizing this instrument are as per the following:

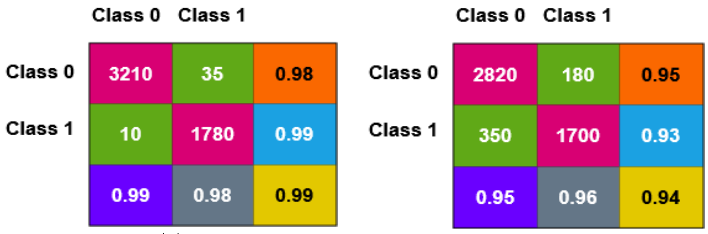
- Demands least measure of investment for preparing and testing the models.
- Limited computational intricacy.
- Expanded discovery proficiency and exactness.
- Guaranteed unwavering quality and versatility.

## 4. RESULTS AND DISCUSSION

This section validates the results of both existing and proposed remote disease detection methodologies by using various evaluation metrics, where the MATLAB/simulation tool is utilized to obtain the results. For this analysis, the diabetics and heart disease detection datasets are used, which includes Hungarian and PIMA (Indian Diabetes Database). The Hungarian dataset comprises 76 attributes, in which some of the essential attributes are listed in Table 2. Figure 4 which has Figure 4(a) and Figure 4(b) shows the confusion matrix of the heart disease dataset with respect to the classes of 0 and 1. Based on the results, it is analyzed that the proposed mechanism accurately predicts the classes with increased TPR. Similarly, the PIMA dataset is obtained from the machine learning repository, which holds 768 number of samples with 9 different feature attributes. Also, it comprises the patient information aging from 21 to 81, and its attribute information are listed in Table 3.

Table 2. Hungarian dataset description

| | Description |
|---|---|
| Age | Young, medium, old or very old |
| Gender | Male and female |
| Maximum Heart Rate (MHR) | Low, medium, high or very high |
| Chest Pain (CPT) | Type 1- typical<br>Type 2 – Atypical<br>Type 3 – Non-anginal<br>Type 4 – Asymptomatic |
| Resting Blood Pressure (RBP) | Low, medium, high or very high |
| Blood sugar (FBS) | BS > 120 mg/dl |
| Serum Cholesterol (SCH) | Low, medium, high or very high |
| Major vessels (VCA) | Major vessels (0 to 3) |
| Thallium scan (TCA) | Normal, fixed defect, and reversible defect |
| Exercise Induced Angina (EIA) | Yes – 1<br>No – 0 |
| Depression Induced by Exercise (OPK) | Low, Risk or Terrible |




(a) (b)

Figure 4. Confusion matrix for Hungarian dataset (a) training and (b) testing

Table 3. PIMA dataset description

| Attributes | Description |
|---|---|
| No of pregnancies | Numerical attribute |
| Blood Pressure (BP) | mmHg |
| Age | Young, middle, old or very old (in terms of years) |
| BMI | Body mass index |
| Insulin | mU/mL |
| Glucose level | Estimated in terms of mg/dL |
| Skin type | Mm |
| Diabetes Pedigree | Yes – 1<br>No – 0 |
| No of pregnancies | Numerical attribute |

Figure 5 and Table 4 looks at the current [33] and proposed arrangement procedures in view of the proportions of accuracy, review, precision, and f-measure. These parameters are extensively used in many detection/classification application systems for analyzing the efficiency of methodologies. Also, the overall performance of classifier is highly depending on the improved values of these measures, which are calculated as follows (21)-(24):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{21}$$

$$Precision = \frac{TP}{TP+FP} \tag{22}$$

$$Recall\ or\ TPR = \frac{TP}{TP+FN} \tag{23}$$

$$F-Measure = \frac{2TP}{2TP+FP+FN} \tag{24}$$

where, TP – true positives, TN – true negatives, FP – false positives, and FN – false negatives. Based on the estimatedresults, it is analyzed that the proposed ULMCSO-PGND outperforms the other approaches with increased precision, recall, accuracy, and f-measure values.

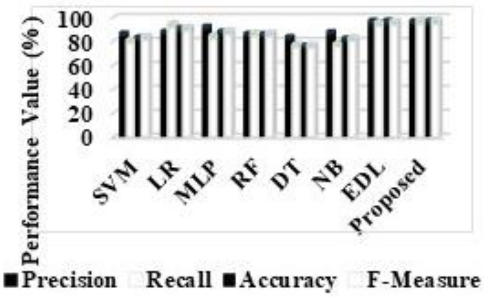



Figure 5. Similar investigation among existing and proposed procedures utilizing Hungarian dataset

Table 4. Accuracy, sensitivity, and specificity analysis using the Hungarian dataset

| Methods | Precision | Recall | Accuracy | F-Measure |
|---|---|---|---|---|
| SVM | 87.5 | 81.5 | 84.4 | 84.5 |
| LR | 89.2 | 95.2 | 92.2 | 92.2 |
| MLP | 93.3 | 85.3 | 89.3 | 89.3 |
| RF | 87.4 | 87.4 | 87.3 | 87.4 |
| DT | 84.6 | 77.7 | 77.6 | 77.6 |
| NB | 88.8 | 78.5 | 83.4 | 83.4 |
| Ensemble DL | 98.2 | 96.4 | 98.5 | 97.2 |
| Proposed | 98 | 98.5 | 98.5 | 98 |

Figure 6 and Table 5 thinks about the root mean squared blunder (RMSE) and mean absolute error (MAE) worth of both existing and proposed arrangement strategies utilizing the Hungarian dataset. These are the error measures calculated as shown in (25), (26):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{25}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \tag{26}$$

here N is the total number of observations, $x_i$ is the actual value, and $\hat{x}_i$ denotes the predicted value. From the obtained results, it is evident that the proposed ULMCSO-PGND technique provides the reduced error outputs, when compared to the other techniques. Because, the normalization and feature selection processes could efficiently reduce the error values by perfectly training the classifier with the optimal features.

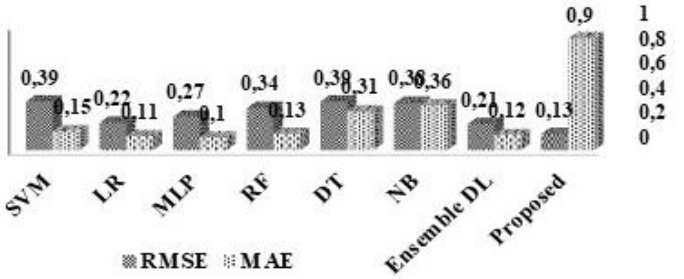



Figure 6. RMSE and MAE

Table 5. RMSE and MAE using the Hungarian dataset

| Methods | RMSE | MAE |
|---|---|---|
| SVM | 0.39 | 0.15 |
| LR | 0.22 | 0.11 |
| MLP | 0.27 | 0.10 |
| RF | 0.34 | 0.13 |
| DT | 0.39 | 0.31 |
| NB | 0.38 | 0.36 |
| Ensemble DL | 0.21 | 0.12 |
| Proposed | 0.13 | 0.9 |

Figure 7 and Table 6 compares the overall accuracy of both existing and proposed classification techniques by using the Hungarian dataset. Typically, the overall accuracy is estimated for analyzing the total efficiency and improved performance of the detection mechanism. According to this evaluation, it is observed that the proposed ULMCSO-PGND technique outperforms the other approaches with increased accuracy. Due to the optimal selection of feature, the training model of classifier is improved, which helps to obtain an increased accuracy value.

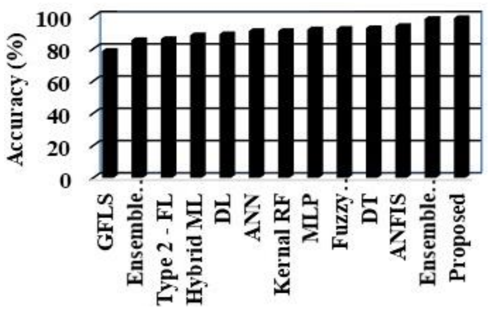



Figure 7. Overall accuracy of existing and proposed classification techniques

Table 6. Overall accuracy analysis using Hungarian dataset

| Methods | Overall accuracy |
|---|---|
| GFLS | 78.7 |
| Ensemble classifier | 85.4 |
| Type 2 - FL | 86 |
| Hybrid ML | 88.4 |
| DL | 89 |
| ANN | 91 |
| Kernal RF | 91 |
| MLP | 92 |
| Fuzzy Diagnosis system | 92.3 |
| DT | 92.8 |
| ANFIS | 94.1 |
| Ensemble DL | 98.5 |
| Proposed | 99 |

Figure 8 and Table 7 glances at the current [34] and proposed portrayal techniques considering the accuracy of disclosures. Where the PIMA dataset is considered for examination. The acquired outcomes portray that the proposed ULMCSO-PSND strategy gives an expanded exactness, when contrasted with different procedures.

Figure 9 and Table 8 looks at the current [35] and proposed arrangement methods in view of the proportions of exactness, accuracy, and review utilizing the PIMA dataset. In the proposed system, a cross breed ULMCSO calculation finds the best ideal arrangement with diminished number of emphases and expanded combination rate, which assists with choosing the most fit set of highlights. Then, these features are used to train the classifier that improves the entire performance of system with increased accuracy, precision, and recall values.

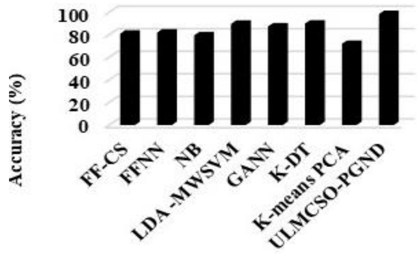



Figure 8. Accuracy analysis using PIMA dataset

Table 7. Comparative analysis based on accuracy using PIMA dataset

| Methods | Accuracy (%) |
|---|---|
| Firefly with cuckoo optimization-based classification | 81 |
| Feed forward NN | 82 |
| NB | 79.56 |
| LDA with MWSVM | 89.74 |
| GA with NN | 87.46 |
| K-means with DT | 90 |
| K-means with PCA | 72 |
| ULMCSO-PGND | 98.5 |

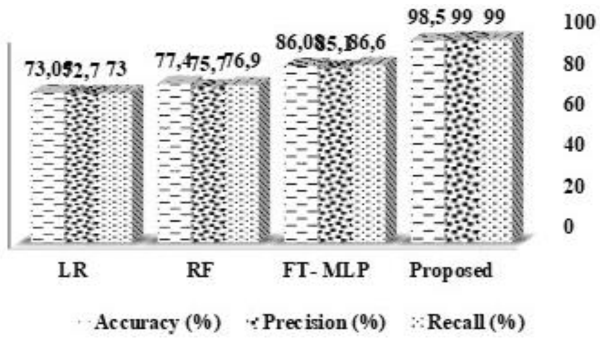



Figure 9. Overall comparative analysis using PIMA dataset

Table 8. Accuracy, precision, and recall analysis using PIMA dataset

| Methods | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| LR | 73.05 | 72.7 | 73 |
| RF | 77.4 | 75.7 | 76.9 |
| Fine-tuned MLP | 86.08 | 85.1 | 86.6 |
| Proposed | 98.5 | 99 | 99 |

## 5. CONCLUSION

This paper presented a new remote health monitoring framework using an advanced AI method for predicting the different types of disease. The main contribution of this paper is to remotely monitor the health status of patients for disease identification, and proper diagnosis. In this framework, the one-time authentication is performed for fusing the medical information obtained from the sensors. Then, the mobile application can directly extract the required data from the cloud platforms, hence the user does not require to enter their data, which is one of the key benefits of this system. For this purpose, the Google Fit cloud service is utilized in this model that maintains the medical information of patients obtained through smart devices. Moreover, the different types of nodes such asdetecting node, visualization node, and prognostic node are used to incorporate the medical information obtained from the different sources. During sickness expectation, the givenclinical datasets are preprocessed and standardized at the underlying state for working on the exhibition of classifier. Then, at that point, the ULMCSO computation is applied to pick the components for decreasing the dimensionality and extending the accuracy of classifier preferably. After enhancement, the PGND characterization component is executed to identify the sickness from the given datasets by utilizing the ideal number of chosen highlights. For approving the outcomes, two famous datasets, for example, PIMA and Hungarian are utilized to assess the presentation of proposed model. Additionally, the acquired outcomes are contrasted and the new best in class approaches as far as exactness, accuracy, review, f-measure, and mistake rate. From the evaluation, it is analyzed the proposed outperforms the other techniques with improved performance values.

## REFERENCES

[1] N. Kalid, A. Zaidan, B. Zaidan, O. H. Salman, M. Hashim, and H. Muzammil, "Based real time remote health monitoring systems: A review on patients prioritization and related big data using body sensors information and communication technology," *Journal of medical systems*, vol. 42, pp. 1-30, 2018, doi: 10.1007/s10916-017-0883-4.

[2] O. S. Albahri *et al.*, "Systematic review of real-time remote health monitoring system in triage and priority-based sensor technology: Taxonomy, open challenges, motivation and recommendations," *Journal of medical systems,* vol. 42, pp. 1-27, 2018, doi: 10.1007/s10916-018-0943-4.

[3] O. A. Shorman, B. A. Shorman, M. Alkhassaweneh, and F. Alkahtani, "A review of internet of medical things (IoMT)–based remote health monitoring through wearable sensors: A case study for diabetic patients," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 20, no. 1, pp. 414-422, 2020, doi: 10.11591/ijeecs.v20.i1.pp414-422.

[4] J. Divakaran *et al.*, "Improved handover authentication in fifth-generation communication networks using fuzzy evolutionary optimisation with nanocore elements in mobile healthcare applications," *Journal of Healthcare Engineering,* vol. 2022, 2022, doi:. 10.1155/2022/2500377.

[5] M. Talal *et al.*, "Smart home-based IoT for real-time and secure remote health monitoring of triage and priority system using body sensors: Multi-driven systematic review," *Journal of medical systems,* vol. 43, pp. 1-34, 2019, doi:. 10.1007/s10916-019-1158-z.

[6] M. Hamim, S. Paul, S. I. Hoque, M. N. Rahman, and I.-A. Baqee, "IoT based remote health monitoring system for patients and elderly people," in *2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST),* 2019, pp. 533-538, doi: 10.1109/ICREST.2019.8644514.

[7] H. Manoharan *et al.*, "A machine learning algorithm for classification of mental tasks," *Computers and Electrical Engineering,* vol. 99, p. 107785, 2022, doi: 10.1016/j.compeleceng.2022.107785.

[8] A. Gondalia, D. Dixit, S. Parashar, V. Raghava, A. Sengupta, and V. R. Sarobin, "IoT-based healthcare monitoring system for war soldiers using machine learning," *Procedia computer science,* vol. 133, pp. 1005-1013, 2018, doi: 10.1016/j.procs.2018.07.075.

[9] L. P. Malasinghe, N. Ramzan, and K. Dahal, "Remote patient monitoring: a comprehensive study," *Journal of Ambient Intelligence and Humanized Computing,* vol. 10, pp. 57-76, 2019, doi: 10.1007/s12652-017-0598-x.

[10] P. N. Ramkumar *et al.*, "Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning–based surveillance platform," *The Journal of Arthroplasty,* vol. 34, pp. 2253-2259, 2019, doi: 10.1016/j.arth.2019.05.021.

[11] S. P. Chatrati *et al.*, "Smart home health monitoring system for predicting type 2 diabetes and hypertension," *Journal of King Saud University-Computer and Information Sciences,* 2020.

[12] S. Vitabile *et al.*, "Medical data processing and analysis for remote health and activities monitoring," in *High-Performance Modelling and Simulation for Big Data Applications*, ed: Springer, Cham, pp. 186-220, 2019, doi: 10.1007/978-3-030-16272-6_7.

[13] L. R. Nair, S. D. Shetty, and S. D. Shetty, "Applying spark based machine learning model on streaming big data for health status prediction," *Computers and Electrical Engineering,* vol. 65, pp. 393-399, 2018, doi: 10.1016/j.compeleceng.2017.03.009.

[14] H. Li, A. Shrestha, H. Heidari, J. L. Kernec, and F. Fioranelli, "A multisensory approach for remote health monitoring of older people," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology,* vol. 2, pp. 102-108, 2018, doi: 10.1109/JERM.2018.2827099.

[15] P. Kaur, R. Kumar, and M. Kumar, "A healthcare monitoring system using random forest and internet of things (IoT)," *Multimedia Tools and Applications,* vol. 78, pp. 19905-19916, 2019, doi: 10.1007/s11042-019-7327-8.

[16] M. Padmaja, S. Shitharth, K. Prasuna, A. Chaturvedi, P. R. Kshirsagar, and A. Vani, "Grow of artificial intelligence to challenge security in IoT application," *Wireless Personal Communications,* pp. 1-17, 2021, doi: 10.1007/s11277-021-08725-4.

[17] P. Verma and S. K. Sood, "Fog assisted-IoT enabled patient health monitoring in smart homes," *IEEE Internet of Things Journal,* vol. 5, pp. 1789-1796, 2018, doi: 10.1109/JIOT.2018.2803201.

[18] I. Azimi, J. Takalo-Mattila, A. Anzanpour, A. M. Rahmani, J.-P. Soininen, and P. Liljeberg, "Empowering healthcare IoT systems with hierarchical edge-based deep learning," in *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies,* 2018, pp. 63-68, doi: 10.1145/3278576.3278597.

[19] A. Mohsin *et al.*, "Real-time remote health monitoring systems using body sensor information and finger vein biometric verification: A multi-layer systematic review," *Journal of medical systems,* vol. 42, pp. 1-36, 2018, doi: 10.1007/s10916-018-1104-5.

[20] M. Devarajan, V. Subramaniyaswamy, V. Vijayakumar, and L. Ravi, "Fog-assisted personalized healthcare-support system for remote patients with diabetes," *Journal of Ambient Intelligence and Humanized Computing,* vol. 10, pp. 3747-3760, 2019, doi: 10.1007/s12652-019-01291-5.

[21] S. R. Hassan, I. Ahmad, S. Ahmad, A. Alfaify, and M. Shafiq, "Remote pain monitoring using fog computing for e-healthcare: An efficient architecture," *Sensors,* vol. 20, p. 6574, 2020, doi: 10.3390/s20226574.

[22] M. Hasan, M. Shahjalal, M. Z. Chowdhury, and Y. M. Jang, "Real-time healthcare data transmission for remote patient monitoring in patch-based hybrid OCC/BLE networks," *Sensors,* vol. 19, p. 1208, 2019, doi: 10.3390/s19051208.

[23] Z. Al-Makhadmeh and A. Tolba, "Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach," *Measurement,* vol. 147, p. 106815, 2019, doi: 10.1016/j.measurement.2019.07.043.

[24] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE access,* vol. 7, pp. 81542-81554, 2019, doi: 10.1109/ACCESS.2019.2923707.

[25] M. R. Nalluri and D. S. Roy, "Hybrid disease diagnosis using multiobjective optimization with evolutionary parameter optimization," *Journal of healthcare engineering,* vol. 2017, 2017, doi: 10.1155/2017/5907264.

[26] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mobile Information Systems,* vol. 2018, 2018, doi: 10.1155/2018/3860146.

[27] K. Uyar and A. İlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks," *Procedia computer science,* vol. 120, pp. 588-593, 2017, doi: 10.1016/j.procs.2017.11.283.

[28] F. Ahmed, "An internet of things (IoT) application for predicting the quantity of future heart attack patients," *International Journal of Computer Applications,* vol. 164, pp. 36-40, 2017, doi: 10.5120/ijca2017913773.

[29] S. Nazari, M. Fallah, H. Kazemipoor, and A. Salehipour, "A fuzzy inference-fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases," *Expert Systems with Applications,* vol. 95, pp. 261-271, 2018, doi: 10.1016/j.eswa.2017.11.001.

[30] T. Vivekanandan and N. C. S. N. Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease," *Computers in biology and medicine,* vol. 90, pp. 125-136, 2017, doi: 10.1016/j.compbiomed.2017.09.011.

[31] L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed, and J. A. Khan, "An automated diagnostic system for heart disease prediction based on ${\chi^{2}} $ statistical model and optimally configured deep neural network," *IEEE Access,* vol. 7, pp. 34938-34945, 2019, doi: 10.1109/ACCESS.2019.2904800.

[32] J. Nourmohammadi-Khiarak, M.-R. Feizi-Derakhshi, K. Behrouzi, S. Mazaheri, Y. Zamani-Harghalani, and R. M. Tayebi, "New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection," *Health and Technology,* vol. 10, pp. 667-678, 2020, doi: 10.1007/s12553-019-00396-3.

[33] F. Ali *et al.*, "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Information Fusion*, vol. 63, pp. 208-222, 2020, doi: 10.1016/j.inffus.2020.06.008.
[34] H. Naz and S. Ahuja, "SMOTE-SMO-based expert system for type II diabetes detection using PIMA dataset," *International Journal of Diabetes in Developing Countries*, pp. 1-9, 2021, doi: 10.1007/s13410-021-00969-x.
[35] U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "machine learning based diabetes classification and prediction for healthcare applications," *Journal of healthcare engineering*, vol. 2021, 2021, doi: 10.1155/2021/9930985.

## BIOGRAPHIES OF AUTHORS

**Adari Ramesh** received his Bachelor degree from Kuvempu University during the year 1998, M. Tech from Jawaharlal Nehru Technological University Hyderabad (JNTUH) during the year 2019. He is working as a architect and research head in Department of R&D, Wipro. He had worked with EPAM solutions Hyderabad, Satyam, Convergys, Conseco Inc, Globarena web Technologies, ElechemSaknyaSastraya PVT Ltd, having more than 22 years of Professional experience which includes Software Industry. His areas of interests are data science, artificial intelligence and mechine learning, IBM common rule engine, databases, deep learning, nural networks. He had diversified domain exposure on products and software development experience as a solutions architect where he executed solutions in various domains like manufacturing, insurance, health care, banking, digital education, telecom billing, life science, auto motive and aero space. He can be contacted at email: adariramesh3@gmail.com.

**Dr. Ceeke Kalappagowda Subbaraya** has completed his Primary and Secondary education at Balehonnur. He did his Bachelor’s Degree from University of Mysore (1977) and obtained his Master Degree from University of Mysore with IV Rank (1980). He did Post Graduate Diploma in Computer Science from University of Mysore in 1983 with I-Rank and obtained Ph.D. degree from UGC-Center for Advanced Studies in Fluid Mechanics, Bangalore University in 1993 with a specialization in Fluid Mechanics. His expertise subjects in applied mathematics, computational methods, computer science. He has published many research papers in National and International Journal/Proceedings. Also he has presented several papers in National and International Conferences and Seminars. He has conducted many National and International Conferences and Seminars. He in the editorial board of National and International journals. He can be contacted at email: registrar@acu.edu.in.

**Dr. Ravi Kumar Guralamata Krishnegowda** received his Bachelor degree from Bangalore University during the year 1996, M. Tech from Karnataka Regional Engineering College Surthkal (NITK) during the year 2000 and Ph.D from Dr MGR University, Chennai. He is working as a Professor and Research Head in Department of Computer Science Engineering, BGSIT. He had worked with iGATE Global solutions Bangalore, Wipro and also has worked with SJBIT as Prof and HOD of Dept of CSE and ISE, having more than 20 years of Professional experience which includes Software Industry and teaching experience. His areas of interests are data warehouse and business intelligence, multimedia, databases, AI, machine learning. He has published and presented papers in national and international conferences and journals. He can be contacted at email: ravikumargk@bgsit.ac.in.