IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 February 2023, accepted 13 March 2023, date of publication 17 March 2023, date of current version 5 April 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3258549*

RESEARCH ARTICLE

# A Diabetes Monitoring System and Health-Medical Service Composition Model in Cloud Environment

**SANTOSH KUMAR SHARMA[1], (Graduate Student Member, IEEE), ABU TAHA ZAMANI[2], AHMED ABDELSALAM[3], DEBENDRA MUDULI[1], AMERAH A. ALABRAH[4], NIKHAT PARVEEN[5], AND SULTAN M. ALANAZI[2]**

[1]Department of Computer Science and Engineering, C.V. Raman Global University, Bhubaneswar, Odisha 752054, India
[2]Department of Computer Science, Northern Border University, Arar 73211, Saudi Arabia
[3]STC's Artificial Intelligence Chair, Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia
[4]Department of Information Systems, College of Computer and Information Sciences, King saud University, Riyadh 11543, Saudi Arabia
[5]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh 522302, India

Corresponding author: Debendra Muduli (muduli.debendra@gmail.com)

This work was supported by the Deanship of Scientific Research at Northern Border University, Arar, Saudi Arabia, under Grant NBU-FFR-2023-0002.

**ABSTRACT** Diabetes is a common chronic illness or absence of sugar in the blood. The early detection of this disease decreases the serious risk factor. Nowadays, Machine Learning based cloud environment acts as a vital role in disease detection. The people who belong to the rural areas are not getting the proper health care treatments. So, this research work proposed an automated eHealth cloud system for detecting diabetes in the earlier stage to decrease the mortality rate and provides health treatment facilities to rural peoples. Extreme Learning Machine (ELM) is a type of Artificial Neural Network (ANN) that has a lot of potential for solving classification challenges. This research work is consisting of several activities like feature normalization, feature selection and classification. We have employed principal component analysis (PCA) for feature selection and extreme learning machine (ELM) for classification. Finally, a cloud computing-based environment with three numbers of virtual machines (vCPU-4, vCPU-8, and vCPU-16), is used for the detection of diabetes. The efficacy of the proposed model has been evaluated with the PIMA dataset in both standalone and cloud environments and achieved 90.57 % accuracy, 82.24 % sensitivity, 73.23 % specificity, and 75.03 % F-1 score with the virtual machine vCPU-16. The experimental results define the proposed model as superior to other state-of-art models with better classification accuracy and less number of features.

**INDEX TERMS** Random forest, K-nearest neighbour, extreme learning machine, principal component analysis, attribute weighted artificial immune system, diabetes mellitus, neural network, extremely randomized trees classifier.

## I. INTRODUCTION

Diabetes (Insulin disorder) is a hormonal illness that affects individuals and is potentially common in modern society. It is a significant issue that can be connected to numerous ailments anywhere in the world. Every year, there are millions more people, including children, who sustain injuries. This illness diagnosis involves routine manual inspection, and automation is a relatively new development in the biomedical industry. The traditional approach predicts diabetes using a single algorithm. The use of a single system to solve complex issues is insufficient and may not be adequate considering the input data for this model. Multiple algorithms are utilized to solve difficult issues. Both homogeneous models and heterogeneous models of these methods are possible. This study uses a stacking ensemble model, a heterogeneous ensemble technique, to identify if people have suffered from diabetes or are non-diabetic. This stacking prediction method can be

The associate editor coordinating the review of this manuscript and approving it for publication was Ziyan Wu.

used to forecast outcomes and evaluate alternative models. Foods for humans are extremely high in portions, calories, fats, and minerals. A person may get diabetes for three main reasons: genetics, environment, and an affluent lifestyle. The most important factor is inherited since family research has shown that children of parents who have type 2 [Muoio and Newgard [1]] diabetes have a greater risk of starting to develop the disease as a result of that person's lifestyle even when their ancestors did not have such diseases. The next important factor is a lifestyle for diabetics who are confirmed by research to have a poor diet and also do regular exercise. The third cause is using complex weight loss techniques. It causes heart problems or renal failure, which eventually results in post-diabetes. Retinopathy, unexpected weight loss, frequent urination, frequent hunger, and push are all signs of acute diabetes. This kind of disease increases the chance of diabetes mellitus nationwide. The importance of this chronic illness in society has grown. It is further broken down into Type I (insulin-dependent diabetes/juvenile onset), Type II (non-insulin-dependent diabetes), Type III (secondary diabetes), and Type IV (adult-onset diabetes) (Gestational DM). According to the International Diabetes Federation (IDF), 463 million people will have diabetes in the United States in 2019, with that number expected to rise to 700 million by 2045. Patients with prediabetes have lower medical costs, a lower chance of death, and less access to quality health care in their area. In 2014, diabetes was estimated to have 422 million. In 2040, the number will reach 642 million [2].

Diabetic patients having without age across all areas of the world and have no preventive precautions to be taken. Type 2-positive patients are always unbalanced in hormones.

Only after the right amount of insulin is produced by beta cells can the insulin raise or lower blood sugar levels to conserve energy in the cells. People today are interested in saving a lot of money. It frequently has several health issues, including consequences like diabetes. One of the self-processes for diabetes detection in humans at an early stage has presented a model. Furthermore, the proposed model was evaluated against other ones that were already in use and outperformed them in terms of accuracy. In 2019, electronic health documents are continuously stored in the healthcare protection system. Due to more access to a maximum quantity of patient data, healthcare offers a new style to hub on expanding their society for controlling and utilizing data mining. Medical databases have gathered huge amounts of health-related data and medical situations. Regrettably, to find out the knowledge, the least technologies have been developed and implemented [3]. Data mining has a significant impact on controlling huge information and also finding interesting facts which can be capable to manage. In the medical field system, data mining methods have been applied in various areas like disease detection and prediction, predictive medicine, coordination system, fraud detection system, and manipulation of the result of some treatments [4], [5], [6]. Medical mining is a vast area of research in mining systems followed to identify and treat problems as well as know the improvements of several chronic diseases. It includes learning from hospital information, data based on health care, and get knowledge from disaster data. In medical mining, inter-related are classified broadly into different groups according to their constraint. Data mining has many applications in different areas such as network authentication, medical data extraction, cloud computing, and so on [7].

Cloud computing services offer a dashboard that can be accessed through a web browser. This makes it easier for the user to use cloud services. Cloud computing makes it easy to combine data on the cloud, which makes it easier to keep medical records up to date [35], [58]. Cloud computing also has a lot of resources that can hold large amounts of biomedical images or speech data [8]. One of the most crucial matters about cloud computing is that services are always available, which can help biomedical systems keep running with a minimum downstream [9]. Creating environments that help people live independently is impossible without cloud computing services [10]. Additionally, in many developing nations, where medical facilities and knowledge are scarce, cloud environments have been utilized to identify patients, the venerable, and the disabled in towns and villages that are far away or hard to reach [11]. Cloud computing also makes it possible to provide important services, like a quick way to find blood and organ donors in an emergency [12]. ELMs have become popular because they can learn quickly and cost even less to run than other types of machines. The following are the most important findings from this investigation. A cloud-based disease detection system for remote user health data monitoring for diabetes is proposed. By analyzing, cloud servers store health information about people. This method can be used to detect and specify a wide range of illnesses. ELM has been utilized to specify disease details to detect diabetes. Classification methods that have been around for a while are compared to the ELM model. The cloud is used to store and process large datasets and to speed up the execution of disease detection processes, which are then compared using both the cloud and a local platform. The model's ability to classify has been improved by tuning the hidden layer nodes of ELM and using "feature selection" to get rid of selected features. We compare the best ELM performance results in both standalone and cloud environments [7], [11].

The remaining article is structured the following: The related works explain in Section II. In Section III Proposed methodology has been discussed. Research materials and methods have been explained in Section IV, Section V contains Experimental results and discussion and at the end, Section VI articulated on conclusion and future work.

## II. RELATED WORKS

In recent past centuries, several Computer Aided Detection techniques have been applied for diabetes detection. Feature selection and classification form the basis for several of the models of the eHealth cloud system. In diabetes classification, dimensional reduction and normalization play a vital part that influencing diabetes detection for a better outcome. Several models based on diabetes detection use machine learning algorithms to find better accuracy. Many scholars are performing experiments for diabetes classification algorithms of machine learning models like naïve Bayes (NB), J48, decision tree (DT), etc. that work better in the detection of different diseases [8], [9]. Esposito et al. [14] have proposed a machine learning concept that is applied to a Decision tree that predicts diabetes classification at a particular age, leading to high accuracy.

Liu et al. [15] have focused that, the main merit of PCA is that keeping back the segregated facts of each feature split is a vital part. The PCA is employed for feature dimensionality reduction and is used in various machine learning applications [52], [54], [10], [11], [13]. Mohanty et al. [16] have suggested a computer-aided design (CAD) model for classifying breast cancer. This model is simulated with different classifiers like KNN, RF, and Naïve Bayes. Kanchan et al. [13] have utilized PCA to find the minimum number of features to enhance the precision of various supervised machine learning algorithms. They have deployed SVM, NB, and DT classifiers in diabetes disease prediction. In the field of diabetes research, Kavakiotis et al. [18] have come up with a new CAD model based on machine learning and data mining. They have used SVM, ANN, and DT classifiers for feature selection filter methods and rapper methods to wrap up the whole feature selection process. Kononenko [19] has presented a new CAD model deployed on machine learning for medical diagnosis. He has used several classifiers, like Nave Bayes, neural networks, and decision trees, in the field of medical diagnosis tasks. Cao et al. [20] have proposed on mental health of college students should be monitored during epidemics. They have also applied for screening, diagnosis, and the assessment of the severity of anxiety disorders. Hence, principal component analysis (PCA) has been used to reduce the dimensionality of Reduction. Based on best-outcome supervised CNN, the authors [21], [23] have provided a weakly supervised object localization method for the classification of diabetes. Vembandasamy et al. [25] have defined the NB formula for CVD detection by examining the standards. Radhimeenakshi [26] has discovered a model with a mean precision of 86.43% for coronary cardiovascular disease recommended by using more as an artificial neural network (ANN), in addition to providing a clinical decision-making framework for coronary health problem characterization. Mohebbi et al. [27] have articulated a deep-learning model for detecting type 2 diabetes. Pham et al. [28] generated a deep-learning model for analyzing medical records to predict trajectories. Rezazadeh et al. [29] proposed a neural network model to achieve an effective training novel optimization algorithm for clinical data analysis. Rao et al. [30] developed a combined classifier for disease diagnosis. Iwendi et al. [31] proposed a model called N-sanitization, which is used to specify the unstructured healthcare datasets for various disease diagnoses. Butt et al. [32] adopted LSTM to predict diabetes in the PIMA dataset. Machine learning models played a vital role in data analysis, particularly in healthcare data analysis. Thus, the existing model has adopted machine learning classifiers for diabetic positive prediction.

Nowadays, extreme learning machine (ELM) plays a vital role in various disease detection However, Traditional ELM has limitations [33], such as the large number of hidden nodes required and the ill-posed problem caused by random weight initialization. Several ELM-based techniques for solving classification and regression problems have been developed in the recent decade. For a single hidden layer feed-forward network (SLFN), Huang et al. [22] suggested ELM as a rapid learning algorithm that overcomes the limitations of a gradient-based learning scheme. Muduli et al. [24] have developed an automated computer-aided design model based on moth flame optimization with ELM for breast cancer classification.

From the literature review, we have observed that most of the computer-aided detection systems are not able to achieve better classification accuracy and are computationally slower, which in turn is not suitable for the real-time environment. Most of the existing CAD models required a large number of features for achieving satisfying results. Almost all CAD models are based on various traditional machine-learning approaches. An appropriate feature selection and classifier have remained a major challenge. Also, we have observed extreme learning machine has been used in various fields with various advantages like fast convergence, skipping the local minima and computationally faster. So, we have employed ELM in our research work as a classifier. The proposed work aims to develop a CAD system for diabetes prediction with high classification accuracy less number of features and low computational overhead using the eHealth cloud computing system.

## III. PROPOSED METHODOLOGY

The suggested system is split into three stages: preprocessing, feature selection, and classification using a cloud-based system. Before the feature reduction stage, the main feature matrices have been normalized so that the mean is zero and the variance is one. The PCA method is employed for feature selection. The extreme learning machine (ELM) has been used to perform the classification task. The overall description of the proposed methodology is described as follows.

### A. PROPOSE CLOUD-BASED DIABETES DIAGNOSIS MODEL

This research work has suggested a diabetes diagnosis system based on the cloud and monitoring health data from remote users to find diabetics disease. The method can be easily adapted for the diagnosis and categorization of different one application: determining whether or not an illness is a diabetes. The cloud environment with ELM is combinedly and utilized for the binary classification problem. The proposed model is trained in the cloud. The experimental model is

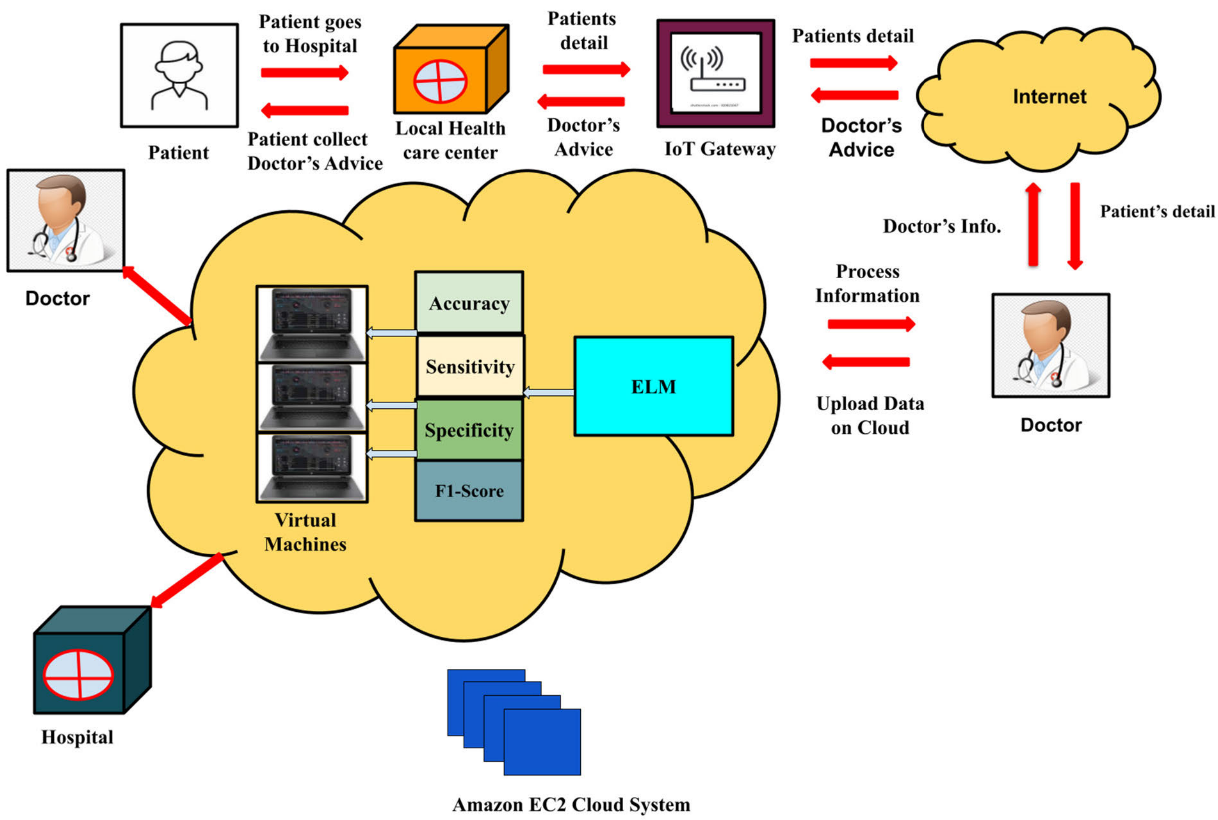


FIGURE 1. Block diagram of the experimental model for the Extreme Learning Machine (ELM) based Cloud environment System.

shown in Figure 1. From the figure it has been observed that he patient has been consulted with a far-off outpatient care facility in interior villages, although the pathologist has collected detailed information from the parent, like fasting blood sugar tests and glucose tolerance tests, then sends the report via the Internet to a doctor; then the doctor uploads the report to the cloud environment for advance treatment.

Two phases of processing make up a typical cloud operation. Attribute selection has been shown to boost the efficiency of machine learning techniques [55], [56] by previous researchers. The initial step is to isolate the important features by employing the PCA technique, after which the unnecessary ones can be filtered out. The proposed model initially learns how to reduce computational complexity. The ELM is then used to put into action for the classification.

### B. FEATURE NORMALIZATION

The proposed system has experimented with PIMA Indian Diabetes dataset [34], [59], [60]. This dataset's fundamental goal by detecting if the patient has diabetic or non-diabetic by focusing on important detection metrics. The PIMA dataset consists of 768 female diabetic patients aged up to 21.

This dataset contains two classes, 268 diabetic patients (having positive) and 500 non-diabetic patients (having negative). Each instance of the dataset has 8 instances of diseases when examining consumer health data stored on cloud servers. In this article, however, we focused largely on pregnancies, glucose, blood pressure, skin thickness, insulin, BMI, diabetes-pedigree-function, and age. The instances have been shown in Table 1.

TABLE 1. Parameters description of PIMA Indian diabetes dataset.

| Sl. No | Parameters | Description |
|---|---|---|
| 1 | pregnancies | few times expecting. |
| 2 | glucose | Blood glucose level after two hours for an oral glucose tolerance test. |
| 3 | blood pressure | pulse rates during diastole. |
| 4 | skin thickness | Skinfold measurements of the triceps. |
| 5 | insulin | 5 mu U/ml insulin accumulation within a week of 2 hours |
| 6 | BMI | The BMI is calculated as weight in kilograms. |
| 7 | Diabetes Pedigree Function | Type 2 diabetes family history. |
| 8 | Age | Age (at least 21 years old) |

The BMI of an adult is measured by dividing their height (meters) and weight(kilograms). An increase in BMI may indicate excessive body fat. If the body mass index (BMI) is less than 18, then we can consider it as underweight, similarly considered as healthy when the range is between 18.5 and 25. It is overweight and obese when the range is between 25.0 - 30.0 and more than 30.0 respectively. The

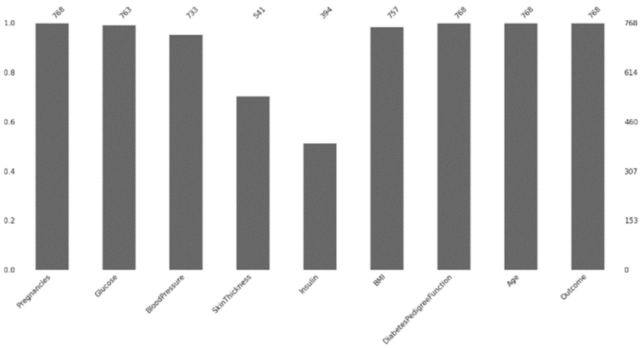


**FIGURE 2.** **Analysis of numerical variables of PIMA Indian dataset.**

Descriptive Statistics

| | count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|
| Pregnancies | 768.00 | 3.85 | 3.37 | 0.00 | 1.00 | 3.00 | 6.00 | 17.00 |
| Glucose | 763.00 | 121.69 | 30.54 | 44.00 | 99.00 | 117.00 | 141.00 | 199.00 |
| BloodPressure | 733.00 | 72.41 | 12.38 | 24.00 | 64.00 | 72.00 | 80.00 | 122.00 |
| SkinThickness | 541.00 | 29.15 | 10.48 | 7.00 | 22.00 | 29.00 | 36.00 | 99.00 |
| Insulin | 394.00 | 155.55 | 118.78 | 14.00 | 76.25 | 125.00 | 190.00 | 846.00 |
| BMI | 757.00 | 32.46 | 6.92 | 18.20 | 27.50 | 32.30 | 36.60 | 67.10 |
| DiabetesPedigreeFunction | 768.00 | 0.47 | 0.33 | 0.08 | 0.24 | 0.37 | 0.63 | 2.42 |
| Age | 768.00 | 33.24 | 11.76 | 21.00 | 24.00 | 29.00 | 41.00 | 81.00 |

**FIGURE 3.** **Descriptive statistics on PIMA Indian dataset.**

numerical analysis of each attribute with the outcome is shown in Figure 2. The descriptive statistics are shown in Figure 3.

The normalization method increases the performance of the classification model [35]. Before applying PCA, the feature matrix (FMa) is normalized having a zero mean and unit variance, and a normalized feature matrix [NM] is obtained. Normalization has been performed to improve the outcomes of the model. The proposed algorithm for normalization is as follows:

**Algorithm 1** Feature Normalization

**Input**: FMa [ 1: T, 1: F]: Feature matrix (T: total number of patients, F: total number of features)

**Output:** NM [1: T, 1: F]: Normalized feature matrix function mean() and std() evaluated the mean and standard deviation of features.

1. Construct a blank matrix NM [1: T, 1: F] and two blanks vectors BV [1, 1: F] and SDM [1,1: F]
2. BV [1,1: F] ←mean (FMa)
3. SDM [1,1: F] ←std (FMa)
4. for i←1 to F do
5. NM [1: T, i] ← (FMa [1: T, i)- BV[ 1, i])/SDM[1, i]
6. end for

## C. FEATURE SELECTION

Feature selection techniques have been proposed to increase the performance of the model by reducing the dimension of the dataset [36], [37]. Here we have taken different feature selection methods like correlation-based feature, PCA, Information Gain Attribute Selection (IGAE), and LDA. The feature selection techniques have been defined as follows:

### 1) CORRELATION-BASED FEATURE SELECTION

Feature selection is a method of choosing relevant instances by detecting diabetes. By Taking an example, if we have bought a building for a specific area, where fn number of attributes related to the building and feature selection process provides by specifying important features from the set of attributes selected, that provide exact results. Features are examined to identify the destination class and Pearson's correlation approach has been used to evaluate the amount of correlation between each instance and instances of the destination part. Nominal instances have been taken into consideration based on the amount and each value affects an indicator [57].

Feature selection elicits a subdivision of important attributes from the selected dataset based on the standard being calculated. A set of attributes are segmented into k number of subparts. Arranging attributes in linear sequence according to their priority. Delicacy can be available neither in the feature vector nor its adjacent feature vector. By avoiding data duplication in two feature vectors, symm- etric uncertainty is used. If two duplicate attributes are available in the dataset, then we must ignore only duplicate attributes, hence each of them will provide us with an equal outcome. List of features in the patients information which is used to examine the healthcare situation of the patients. The algorithm's result is mostly based on feature selection. Best feature selections that are important for the detection of particular attributes but which have no duplication. Coordination between two features is allotted either classical method of linear correlation. In this classical model of liner analysis correlation, for each group pair (t,u), calculate the following result:

$$\mathrm{cr}=\frac{\sum_{\mathrm{g}=0}^{\mathrm{fg}}\left(\mathrm{t_g}-\bar{\mathrm{t}}_{\mathrm{g}}\right)\left(\mathrm{u_g}-\bar{\mathrm{u}}_{\mathrm{g}}\right)}{\sum_{\mathrm{g}=0}^{\mathrm{fg}}\left(\mathrm{t_g}-\bar{\mathrm{t}}_{\mathrm{g}}\right)^{\wedge}2\sqrt{\sum_{\mathrm{g}=0}^{\mathrm{fn}}\left(\mathrm{u_g}-\bar{\mathrm{u}}_{\mathrm{g}}\right)^{\wedge}2}} \tag{1}$$

Here, cr is the coefficient of linear correlation, tg – the average, t and ug is the mean of u. The coefficient of the correlation range between −1 and +1. when the value of the coefficient is zero, then variables t & u are taken as class variables. We create variable t as entropy.

$$He=-\sum_{g=0}^{g}V(t_g)log_2(V(t_g)) \tag{2}$$

The conditional entropy of t provided another variable u can be evaluated as follows:

$$He(\frac{t}{u})=-\sum_{h=0}^{g}V\,(t_h)\sum_{g=0}^{fg}V(\frac{t_g}{u_h})(log_2(V\left(\frac{t_g}{u_h}\right))) \tag{3}$$

Here $\mathrm{V(t_g)}$ is the probability of all values of g and $\mathrm{V}(\frac{tg}{fh})$ is after the probability of the given the value of u. We can create the use of symmetric uncertainty given by last equation and also find out the correlation between the attributes:

$$SUU\ (t,u) = 2[\frac{IGG(\frac{t}{u})}{He\ (t) + He(u)}] \tag{4}$$

#### 2) PRINCIPAL COMPONENT ANALYSIS

The most important feature selection techniques decrease the dimensionality of the attribute set. It is also a procedure based on orthogonal linear transformation in which information can be transmitted in a newly organized manner to specify the principal component [38], i.e., the highest variance, 2nd is correlated with greater variance, and so on. Our proposed dataset has m1× n1 dimensions, with a 0 empirical mean for each column. The empirical mean is the average mean of each column that has been set to zero, and attributes specify a specific attribute from the attribute set, while rows represent exploratory occurrences.

Orthogonal linear transformation [36] is numerically specified by set of countable boundaries that is m1 of n1-dimensional vectors i.e., the coefficient can be evaluated as:

$$\mathrm{Cc(k1)} = (\mathrm{Cc1} \ldots\ldots \mathrm{Cg})\mathrm{k1} \tag{5}$$

Each tuple vector is planned to achieve the new vector of the principal component and is specified by the formula: $tt(kk(g)) = t_g$

$$\mathrm{Cc}(k1) \text{ for } \mathrm{g} = 1, 2, \ldots\ldots\ldots\mathrm{W} \text{ and } \mathrm{k1} = 1, 2, 3 \ldots\ldots \mathrm{v} \tag{6}$$

For evaluation of PC is:

Avoid the balanced element and include the remaining dataset as dd-dimensional.

1. The average of each attribute or feature of the dataset is evaluated.
2. The covariance vector of the entire dataset is evaluated.
3. The eigenvectors and eigenvalues are computed.
4. Eigenvectors are organized in ascending or descending order of eigenvalues, and k1 eigenvectors with the highest eigenvalues are used to build a dd × k1-dimensional vector.
5. The previous calculation vector is implemented for easy transformation into the recent subspace.

#### 3) INFORMATION GAIN ATTRIBUTE SELECTION

It is one of the most vital feature selections that can evaluate the list of attributes about the dependent attributes from which an attribute can select. There is no data available for attributes that cannot be linked to each other [39]. Based on greater information gain entropy, features are ranked chronologically. Entropy can be evaluated by the list of data provided by attributes. Entropy can decrease information gain is one of the major challenges. Entropy can be evaluated as:

$$Ee(Ss) = \sum^{Cc} -V_g log_2 V_g \tag{7}$$

Here, V is the ratio of attributes that have been taken for the dependent variable(class). When the purity level is low then entropy is greater as result [54]. The different steps for information gain calculation are as follows:

1. The entropy has been evaluated for each branch.
2. Before evaluating the entropy, the dataset has been split into different attributes.
3. By adding the entropy of the branch proportionally, the total entropy of the split is computed.
4. The outcome value is deducted from the entropy and provides having final information gain result.

#### 4) LINEAR DISCRIMINANT ANALYSIS (LDA)

It is one of the most vital dimensionality reduction techniques used to solve binary classification problems. It is an algorithm used for predictive model problems also. It is also used to normalize the dataset by avoiding outliers. LDA [40] is a helpful tool for separating feature variables between healthy individuals and updating patients. It is used to decrease the features using the supervised learning method to find the out base vector. The base vector can be specified on the subspace of LDA by simple linear projection. LDA is the interface between the product of the base vector and its related data. In this scheme, a new dimension is extracted which separates the means of the projected class as high and low in projected variance. Fisher method can be classified as a formula:

$$li = \frac{\pi_2 - \pi_1}{\gamma_2 - \gamma_1} \tag{8}$$

Here, $\pi_2$, $\pi_1$ are mean vectors of two dependent variables an $\gamma_2$, $\gamma_1$ are respective variances.

- In phase one, it is used to removable between different classes.
- In the second phase, Class variance is the distance between the mean and sample of each class.
- In the third phase, reducing class variance and increasing low-dimensional space has been designed

### D. EXTREME LEARNING MACHINE (ELM)

In the era of artificial intelligence, ELM [35], [38], [42], [43] refers to a feed-forward neural network architecture that can have one or more layers of hidden nodes and is typically employed for tasks such as classification, regression, clustering, tiny estimate, compression, and pattern learning [24]. Here, the parameters of hidden nodes, like their weights and biases, don't need to be changed. Parameters of hidden nodes, on the other hand, can be assigned arbitrarily and never altered or inherited from their descendants. These models outperform backpropagation-trained networks in terms of learning speed. In backpropagation, the gradients are calculated by iteratively propagating from the network's output to its inputs, which is the most common learning process used in feed-forward neural networks. Still, backpropagation has a number of drawbacks to set the weights and biases after each iteration during training typically takes a long period. While this model strives for precision, it gradually gets less accurate

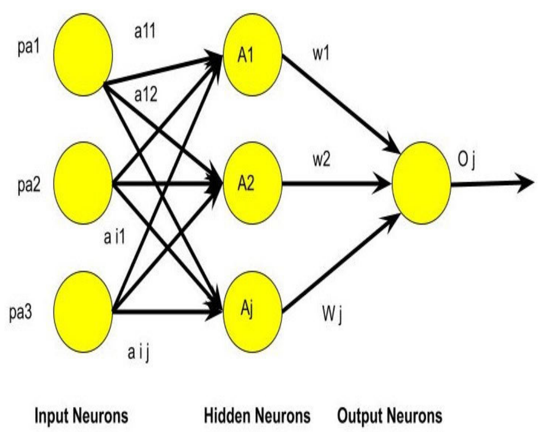


**FIGURE 4. Overall work flow of extreme learning machine.**

since it ignores the weight's magnitude. The effectiveness of the backpropagation learning process is also impacted by the existence of local minima [41], [57]. The ELM network makes it easy to make adjustments to weights and biases. It prioritizes not just obtaining the smallest possible training error but also the lightest possible weight standards, both of which contribute to the model's overall effectiveness. Using these straightforward alternates, we are able to solve the problem of being trapped in local minima and move fast to the global minima. A schematic diagram of ELM is shown in Figure 4.

For AR arbitrary samples $(pp_i, tt_i)$, where $pp_i = [pp_{i1}, pp_{i2} \ldots\ldots pp_{in}]^T \in QQ^n$ and $t_i = [t_{t1i}, t_{t1i} \ldots\ldots t_{in}] \in QQ^m$, the common single-hidden layer feed-forward neural network (SLFNs) with activation function $f(.)$ and GG is the hidden nodes that can be written as

$$\sum_{j=1}^{GG} \omega_i f_i(PP)_i = \sum_{j=1}^{GG} \omega_i f(aa_i \times pp_j \times cc_i) = OO_j, (j = 1, 2 \ldots\ldots HH) \tag{9}$$

Here, $aa_i = [aa_{i1}, aa_{i2} \ldots\ldots aa]^T$ is the weight vector linking $i^{th}$ hidden node and input node $\omega w_i = [\omega w_{i9}, \omega w_i \ldots\ldots \omega w_i]^T$ is the weight vector connecting $i^{th}$ hidden node to the outcome node, $CC_j$ is the threshold of the hidden node, and $OO_j = [OO_{j1}, OO_{j2}, \ldots\ldots OO_{jm,}]^T$ having $j^{th}$ hidden vectors of SLFNs. Standard SLNs with GG hidden nodes and activation function f(.) can finding these HH illustration with zero error, which means that $\sum_{j=1}^{GG} ||OO_j - tt_j|| = 0$ and that there exist $\omega w_i$, $CC_j$ and $OO_j$ such that

$$\sum_{j=1}^{GG} \omega\omega_i f(aa_i \times yy_j \times cc_i) = tt_i (j = 1, 2 \ldots\ldots HH) \tag{10}$$

The all summarization based on the equation that is

$$MM\omega\omega = TT, \tag{11}$$

where, (12)–(14), as shown at the bottom of the page, where MM is called an output matrix of hidden layer and the $k^{th}$ column of MM is the output of the $k^{th}$ hidden node according to input $yy_1, yy_2 \ldots\ldots yy_{HH}$. The outcome of the linear system is

$$\omega\omega = MM^{-1}TT \tag{15}$$

where, $MM^{-1}$ is the Moore-Penrose generalized inverse of matrix MM.

The output function of ELM can be specified as

$$gg(yy) = pp(yy)\,\omega\omega = pp(yy)MM^{-1}TT \tag{16}$$

In proposed extreme learning machine has based on three training key parameters. The training set $KK = \{(yy_i, tt_j) | yy_j \in QQ^{nn}, tt_{jj} \in QQ^{n,}, tt_{jj} \in QQ^m j = 1 \ldots\ldots HH\}$; Then the hidden node of resultant function f $(aa_i, cc_i, yy_j)$ & the number of hidden nodes is GG. Once the appropriate settings have been selected, ELM training can begin. ELM begins by producing parameters for the GG pair of hidden nodes

$$MM(aa_1, aa_2 \ldots\ldots aa_{GG}, cc_1, cc_2 \ldots\ldots cc_{GG}, yy_1, yy_2 \ldots\ldots yy_{GG}) = \begin{bmatrix} f(aa_1 \times yy_1 + cc_1) f(aa_{GG} \times yy_1 + cc_{GG}) \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ f(aa_1 \times yy_{HH} + cc_1) \ldots f(aa_{GG} \times yy_{HH} + cc_{GG}) \end{bmatrix}_{H \times G} \tag{12}$$

$$\omega\omega = \begin{pmatrix} \omega\omega_!^{TT} \\ \vdots \\ \vdots \\ \omega\omega_{NN}^{TT} \end{pmatrix}_{GG \times nn} \tag{13}$$

$$TT = \begin{pmatrix} ii!_{!1}^{TT} \\ \vdots \\ \vdots \\ ii_{NN}^{TT} \end{pmatrix}_{GG \times nn} \tag{14}$$

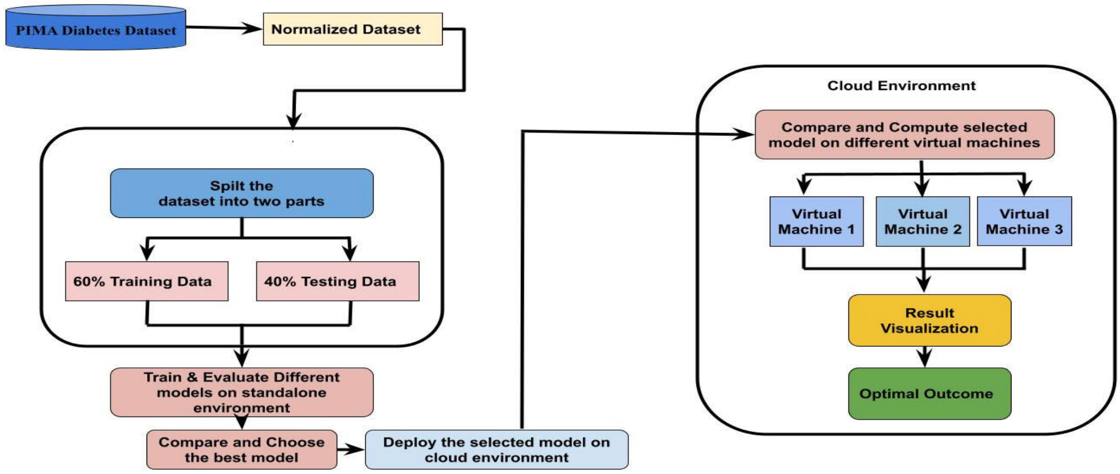


**FIGURE 5.** Work flow diagram of proposed model.

at random ($aa_i$, $cc_i$). Then, using Equation (12), the output matrix M is constructed based on the input and randomly generated parameters. Then, we input those values into Equation (15) to get our weighted output vector $\omega\omega$. After the training phase, Equation (16) can be used to predict the classification result of test data tuples. The steps necessary to train an ELM are as follows. We have given an activation function f (xx), a number of hidden neurons N, and a training set A = f($aa_i$, $dd_i$)| $aa_i$ $\in\in$ $X_n$, $dd_i$ $\in$ $X_n$, i= 1,.....NN},First, a random distribution is used to determine the input wi weights and bias bi weights.

Second, the hidden layer's output matrix MM is calculated.

Then, determine the final mass $\omega\omega$ by using the formula:

$$\omega\omega = MM \times TT \tag{17}$$

Here, MM and TT are used by Equations (11) and (12).

## IV. RESEARCH MATERIALS AND METHODS

The research work has been simulated under two environments like standalone and cloud. The normalized dataset has been divided into two units like training and testing. The PCA method is utilized for feature selection and ELM is used for the classification task. The experimental classification result of ELM has been compared with different traditional machine learning classifiers like K-nearest neighbours [17], [44], Naive Bayes [24], Perceptron network [55], and Support Vector Machine [57]. In the cloud environment, we have used three virtual machines. The overall proposed work is shown in Figure 5. The overall configuration of the standalone and cloud environment is described in further sections.

### A. CLOUD ENVIRONMENT

Platform-as-a-service (PaaS) has been deployed on Amazon Elastic Compute Cloud (EC2) to deploy the ELM model, which has been then compared to the standalone system.

**TABLE 2.** Parameter list.

| Classifiers | Parameters | values |
|---|---|---|
| KNN | Nearest Neighbours | 3 |
| | Searching Algorithm | Euclidean distance |
| | Number of folds | 5 |
| | Batch size | 100 |
| MLP | Initial learning rate | 0.001 |
| | Momentum | 0.9 |
| | Hidden layers | 20 |
| RF | Maximum depth of the tree | 0 |
| | Number of the trees | 4 |
| | Size of each Bag | 52 |
| DT | Confidence factor | 0.10 |
| | Maximum Dept0 | |
| | Maximum size of the tree | 100 |
| | Minimum number of objects | 1 |
| | Unpruned | False |
| ELM | Population size | 50 |
| | Activation Function | Sigmoid |
| | Number of hidden nodes | 300 |

Moreover, the proposed model has applications on image processing, and migrating them to the cloud was done primarily to cut down on latency and improve accuracy. The models can be accessed from anywhere at any time after being moved to the cloud. The Linux operating system has been utilized for the creation of every virtual machine that has been used in the cloud environment [41], [53]. Hence, the results from both environments have been compared. This system has been implemented on the cloud environment provided by Amazon EC2. By using Ubuntu is the operating system of choice for cloud-based virtual machines, with additional configurations like RAM, HDD space, and CPU count being customizable.

### B. STANDALONE ENVIRONMENT

The standalone system has been proposed for an experimental setup like 8 GB RAM, 11th Gen Intel(R) Core (TM) i5-1135G7 @ 2.40GHz processor with clock speed 2.40 GHz. And HDD 1 TB. In this environment, various classification model like python3.1., and PyCharm IDE has been used.

**TABLE 3.** Estimation of matrix for extreme learning machine (ELM) with distinct hidden layer nodes in standalone system.

| ELM + Standalone | 50 | 100 | 150 | 200 | 250 | 300 |
|---|---|---|---|---|---|---|
| Accuracy | 85.79 | 84.58 | 85.57 | 86.84 | **87.88** | 84.46 |
| Sensitivity | 90.43 | 88.75 | 90.97 | 91.12 | 92.12 | 79.02 |
| Specificity | 81.19 | 79.35 | 78.48 | 79.21 | 76.21 | 71.05 |
| F1-Score | 85.79 | 83.75 | 83.91 | 83.36 | 71.24 | 76.48 |

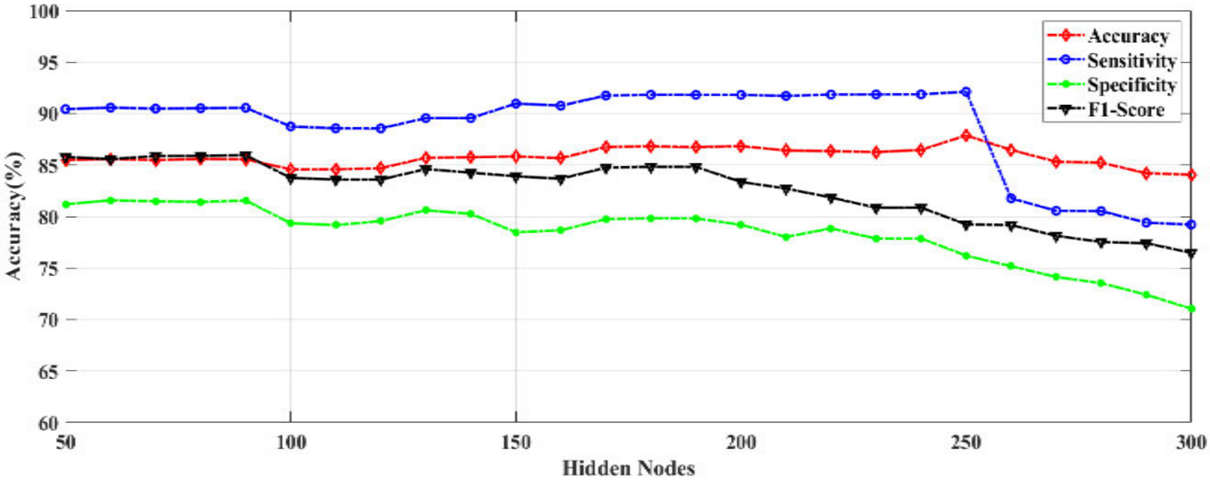


**FIGURE 6.** Diagrammatic display of the model's efficacy (standalone environment).

### C. DATA COLLECTION

In this research work, we have considered the PIMA dataset for experimental purposes. The dataset has been normalized according to Algorithm 1. The PCA method is used for feature selection and considered 5 optimal features Glucose, Insulin, Age, BMI, and Diabetes Pedigree Function.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

A simulation has been validated in local machine as well in the cloud environment with different measuring parameters like sensitivity, specificity, F1 score and accuracy. The parameters are as follows:

- The sensitivity (true positive rate) calculates the number of diabetic patients in PIMA dataset truly specified by the total number of diabetes patients are abnormal data.

$$Sencitivity = \frac{TP}{TP + FN}$$

- Specificity (true negative rate) which can be calculate by the number of patients having not affected by diabetes with total number of non-diabetic patients in PIMA dataset.

$$Specificity = \frac{TN}{TN + FP}$$

- Accuracy is the total number patients of PIMA datasets that are correctly classified.

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

- F1 score is the weighted average of both sensitivity and specificity that are correctly classified.

$$F1Score = 2 \times \frac{Sencitivity \times Specificity}{Sencitivity + Specificity}$$

Here,

True Positive (TP): An individual having predicted diabetic and the actually the person also affected by diabetic.
True Negative (TN): A person having predicted non-diabetic and the real is the person having diabetic.
False Negative (FN): A person having predicted non-diabetic but actually, the person is diabetic
False Positive (FP): A person having predicted diabetic but actually, the person is non-diabetic.

**TABLE 4.** Classification of the proposed model with feature selection methods.

| Classifier | No. of Features | Sensitivity | Specificity | F1 Score | ACC (%) |
|---|---|---|---|---|---|
| LDA+KNN | 6 | 78.02 | 67.21 | 65.21 | 75.00 |
| LDA+NB | 6 | 72.63 | 65.42 | 68.08 | 69.22 |
| LDA+MLP | 6 | 80.00 | 74.33 | 76.54 | 77.71 |
| LDA+RF | 6 | 79.10 | 74.01 | 73.22 | 77.44 |
| LDA+DT | 6 | 82.12 | 75.01 | 75.76 | 78.90 |
| LDA+SVM | 6 | 86.31 | 77.21 | 80.21 | 81.34 |
| LDA+ELM | 6 | 88.19 | 72.08 | 81.33 | 82.53 |
| SAE+KNN | 6 | 89.02 | 72.02 | 81.66 | 82.85 |
| SAE+NB | | 74.63 | 67.68 | 69.45 | 70.77 |
| SAE+MLP | 6 | 84.05 | 77.12 | 78.57 | 81.68 |
| SAE+RF | 6 | 84.06 | 72.11 | 76.23 | 78.83 |
| SAE+DT | 6 | 85.04 | 72.02 | 75.32 | 79.74 |
| SAE+SVM | | 87.32 | 79.04 | 81.32 | 82.77 |
| SAE+ELM | 6 | 87.04 | 76.31 | 82.55 | 83.00 |
| IGAE+KNN | 7 | 84.00 | 74.23 | 80.11 | 80.42 |
| IGAE+NB | 7 | 78.21 | 69.65 | 72.44 | 73.86 |
| IGAE+MLP | 7 | 86.05 | 74.60 | 82.01 | 82.71 |
| IGAE+RF | 7 | 84.31 | 74.11 | 79.32 | 80.93 |
| IGAE+DT | 7 | 82.12 | 71.18 | 75.42 | 77.65 |
| IGAE+SVM | | 87.63 | 79.67 | 83.02 | 84.56 |
| IGAE+ELM | 7 | 86.07 | 74.21 | 80.21 | 82.87 |
| PCA+KNN | 5 | 87.18 | 74.22 | 81.32 | 84.19 |
| PCA+NB | 5 | 80.96 | 74.98 | 72.33 | 78.76 |
| PCA+MLP | 5 | 78.08 | 62.34 | 72.01 | 73.30 |
| PCA+RF | 5 | 80.21 | 73.02 | 76.21 | 77.08 |
| PCA+DT | 5 | 77.26 | 68.15 | 71.23 | 72.52 |
| PCA+SVM | 5 | 89.60 | 81.60 | 84.60 | 86.94 |
| **PCA+ELM** | **5** | **91.12** | **85.21** | **86.24** | **87.88** |

**TABLE 5.** The effectiveness of several traditional models compared with a standalone environment.

| Classifiers | KNN | NB | MLP | RF | DT | SVM | ELM |
|---|---|---|---|---|---|---|---|
| Accuracy | 84.60 | 78.76 | 73.30 | 77.08 | 72.52 | 86.94 | **87.88** |
| Sensitivity | 89.71 | 89.98 | 62.34 | 70.02 | 62.15 | 88.6 | 91.12 |
| Specificity | 88.07 | 75.96 | 64.03 | 69.21 | 56.26 | 88.6 | 85.21 |
| F1-Score | 87.18 | 82.33 | 72.01 | 76.21 | 71.23 | 88.6 | 86.24 |

### *A. ANALYSIS OF PERFORMANCE IN A STANDALONE ENVIRONMENT*

Here we have experimented with the proposed model in both the local and cloud contexts and compared the results. The overall experimental work has been described below. The proposed model has been simulated with different classifiers like k-nearest neighbours (KNN), naïve Bayes (NB), multi-layer perceptron (MLP), random forest (RF), decision

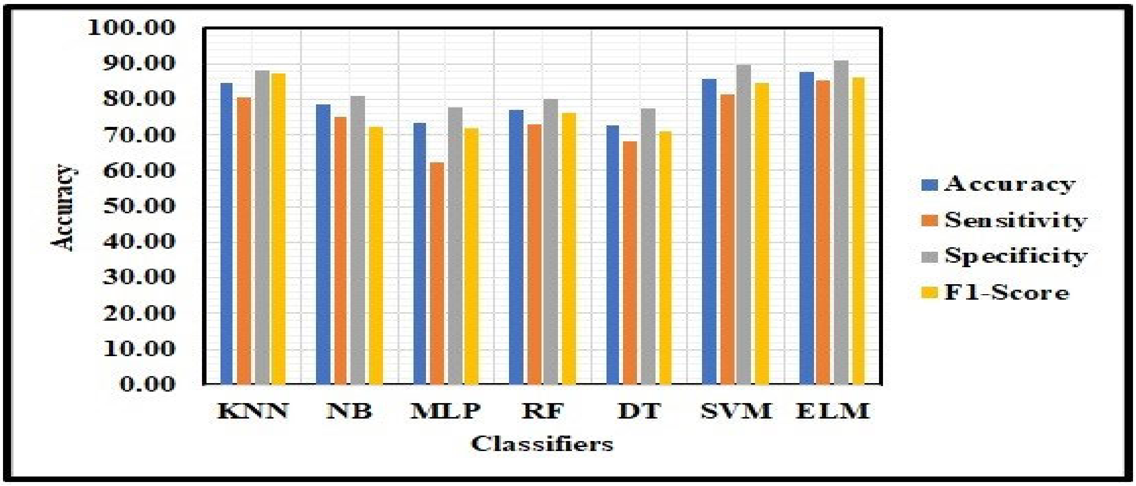


**FIGURE 7.** Diagrammatic display of the several machine learning models compared in a standalone environment.

**TABLE 6.** Analysis of the performance of ELM-based cloud environment with varying numbers of hidden layer nodes. The most precise outcome is highlighted in bold. vCPU-4 16 GB RAM.

| vCPU-4 16 GB RAM. | ELM(50) | ELM(100) | ELM(150) | ELM(200) | ELM(250) | ELM(300) |
|---|---|---|---|---|---|---|
| Accuracy | 84.43 | 85.63 | 86.73 | 87.96 | **88.92** | 84.34 |
| Sensitivity | 89.48 | 86.69 | 88.17 | 89.21 | 82.42 | 86.14 |
| Specificity | 80.38 | 78.61 | 78.11 | 78.57 | 82.32 | 73.02 |
| F1-Score | 84.74 | 82.51 | 82.24 | 82.44 | 76.34 | 83.03 |

**TABLE 7.** Analysis of the performance of ELM based cloud environment with varying numbers of hidden layer nodes. The most precise result is highlighted in bold. vCPU-8 32 GB RAM.

| vCPU-8 32 GB RAM. | ELM(50) | ELM(100) | ELM(150) | ELM(200) | ELM(250) | ELM(300) |
|---|---|---|---|---|---|---|
| Accuracy | 85.19 | 86.13 | 87.66 | 88.66 | **89.88** | 84.58 |
| Sensitivity | 91.71 | 91.71 | 90.01 | 91.81 | 89.92 | 78.02 |
| Specificity | 78.54 | 75.61 | 73.92 | 73.93 | 72.93 | 71.02 |
| F1-Score | 84.34 | 82.81 | 88.24 | 79.44 | 75.54 | 78.33 |

tree (DT), support vector machine (SVM) and extreme learning machine (ELM). The parameters of the different classifiers have been listed in Table 2.

### B. COMPARATIVE STUDY OF ELM'S EFFICACY USING SEVERAL HIDDEN NODE CONFIGURATIONS

The ELM-based proposed model experimented with different number of hidden nodes to boost the proposed model's accuracy in both standalone and cloud environ-ments. During the experiment work, we simulated the model by changing the hidden nodes from 1 to 300. In the standalone environment, we have observed that the model gives various classification accuracy results like 85.79%, 84.58%, 85.75%, 86.8 %, 87.88% and 84.46% with corresponding hidden nodes 50, 100, 150, 200, 250 and 300. Also, we have seen the model provides better accuracy 87.88% with 250 hidden nodes. The different simulated results have been listed in Table 3 and shown in Figure 6.

### C. ANALYSIS OF ELM's EFFICIENCY COMPARED TO OTHER CLASSIFICATION METHODS IN STANDALONE ENVIRONMENT

The ELM classifier, trained with 300 nodes, serves as a benchmark against which the performance of other classifiers

**TABLE 8.** Analysis of the performance of ELM in a cloud computing environment with varying numbers of hidden layer nodes. The most precise values are highlighted in bold. vCPU-16 64 GB RAM.

| vCPU-16 64 GB RAM. | ELM(50) | ELM(100) | ELM(150) | ELM(200) | ELM(250) | ELM(300) |
|---|---|---|---|---|---|---|
| Accuracy | 84.24 | 86.63 | 88.38 | 89.56 | **90.57** | 87.28 |
| Sensitivity | 91.94 | 88.62 | 88.54 | 87.24 | 82.24 | 81.54 |
| Specificity | 82.24 | 80.65 | 72.95 | 73.96 | 73.23 | 71.82 |
| F1-Score | 87.34 | 83.21 | 78.64 | 74.44 | 75.74 | 75.03 |

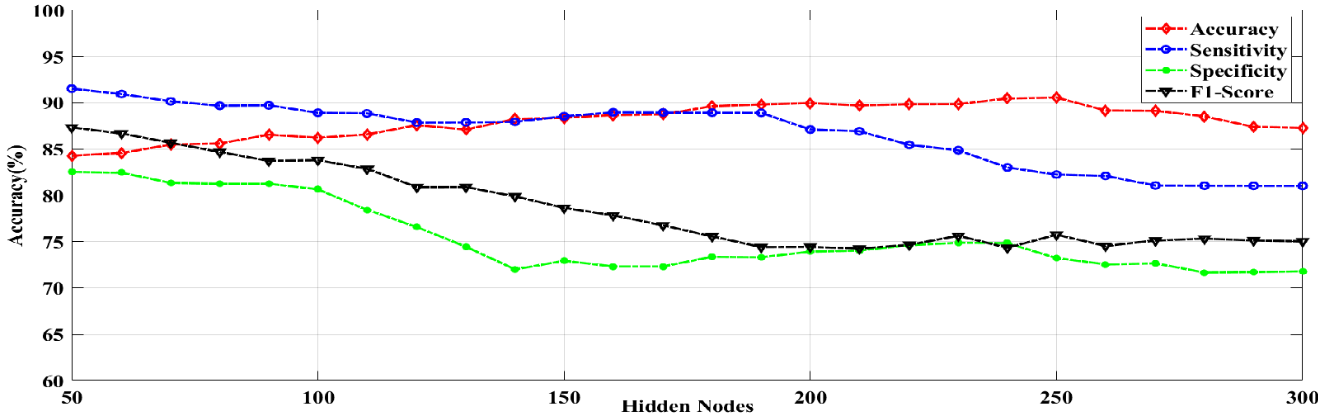


**FIGURE 8.** Visualized representation of ELM model performance (Cloud Computing Environment) on vCPU-4 –16 GB RAM.

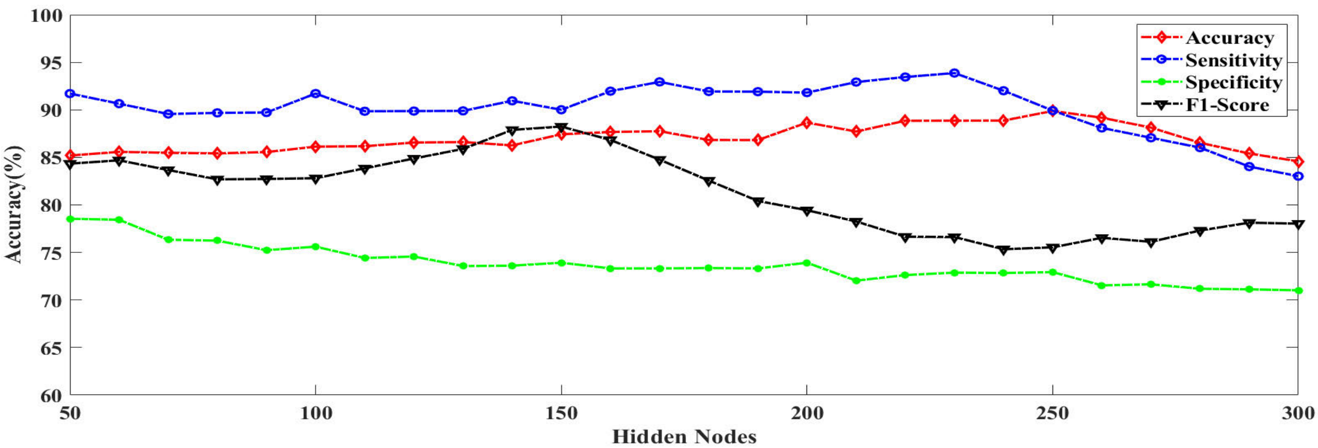


**FIGURE 9.** Graphical representation of ELM model performance (Cloud Computing Environment) on vCPU- 8 –32 GB RAM.

is measured. Keep in mind that only 40% of the dataset's tuples are utilised for testing, while 60% are used for training. We have compared the ELM results with various traditional machine learning classifiers like KNN, NB, MLP, RF, DT and SVM. It has been observed that the DT gives less accuracy as compared to other classifiers. The compared result has been shown in Table 4. with the number of features. From the table, it has been observed that the ELM has achieved better accuracy 87.88% with 5 features. The simulation results of PCA with different classifier has been shown in Table 5 and the result graph in Figure 7.

### D. CLOUD-BASED PERFORMANCE EVALUATION (Amazon EC2)

The proposed model simulated on different hidden nodes in a standalone system, we have been examined on ELM-based

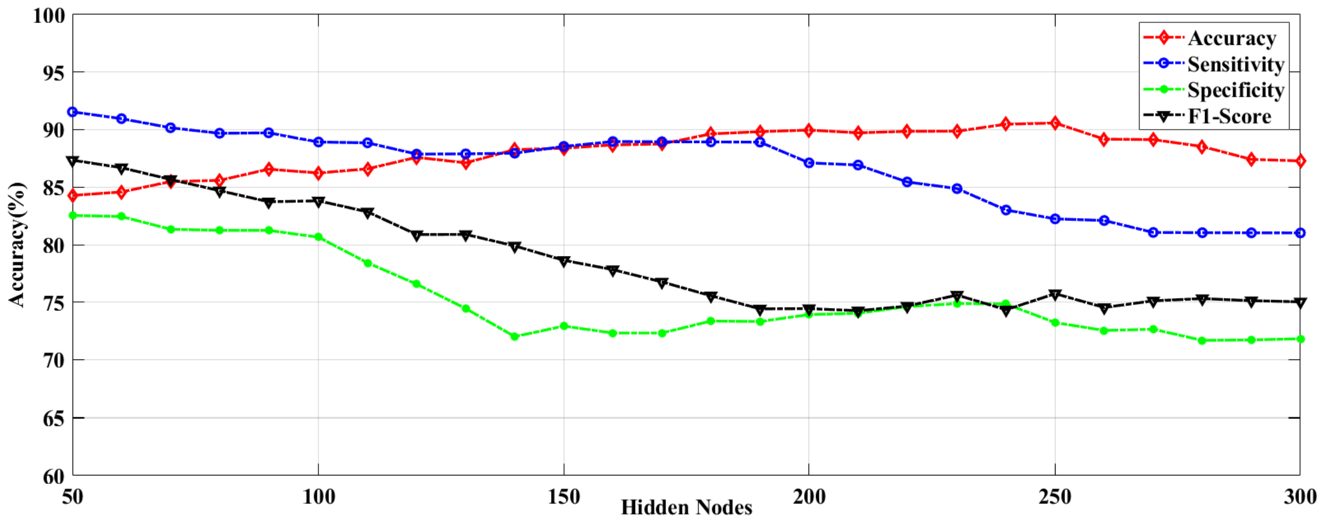


**FIGURE 10.** Graphical representation of ELM model performance (Cloud Computing Environment) on vCPU-16 64 GB RAM.

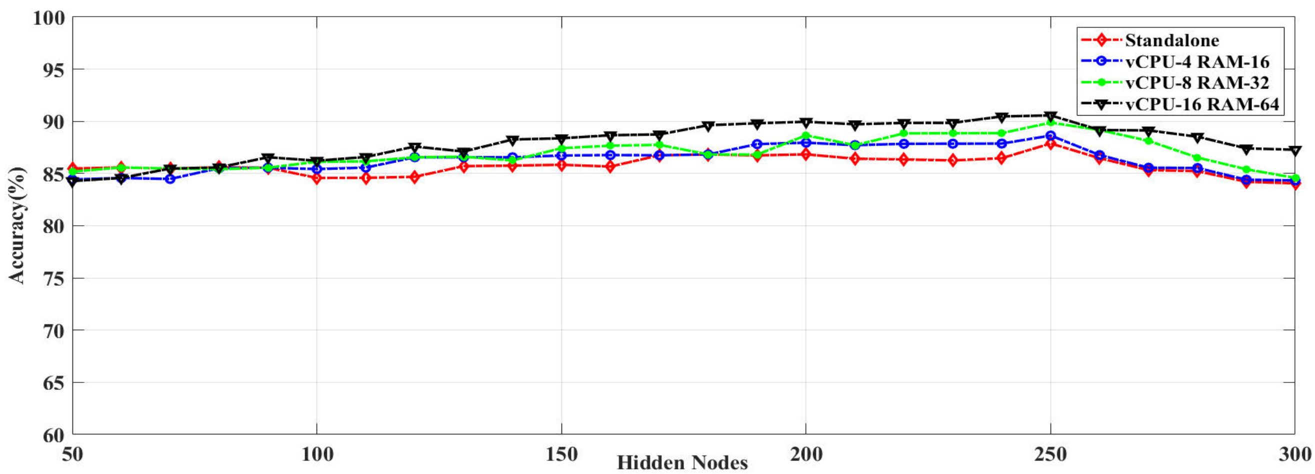


**FIGURE 11.** A comparison of how accurate the results are in a standalone environment and in the cloud.

cloud environment performed better outcomes than other traditional models. So, ELM has been proposed on virtual cloud environment with different hardware specifications. he effectiveness of the ELM in the standalone environment has been found to be improved by altering the hidden layer node.

In the experiment, we considered the number of hidden layers from 1 to 300. In the cloud environment, we have tested the same model with three cloud computing virtual machines with different configurations like vCPU-4 64 GB RAM, vCPU-8 32 GB RAM, and vCPU-16 64 GB RAM. The proposed model has been simulated with three virtual machines with the same number of hidden nodes. The simulated results have been noted on Table 6, Table 7 and Table 8. From the experimental result, it has been discovered that the virtual machine vCPU-16 64 GB RAM has obtained the best accuracy of 90.57% with 250 hidden nodes as compared to other virtual machines. Finally, the results have been visualized in Figure 8, Figure 9, and Figure 10.

### *E. COMPARISON OF ELM ON THE CLOUD ENVIRONMENTS WITH STANDALONE CLOUD ENVIRONMENT*

In this section the performance of the proposed model in standalone and cloud environment has been explained. During the experiment the models have been simulated on the basis of different hidden nodes ranging from 1 to 300. From the experiment we have seen that the model provides better classification accuracy 90.57 % with virtual machine vCPU-16 60 GB RAM in cloud environment as compared to standalone environment.

During execution, the proposed model takes 4.34s in standalone environment and 3.25s in cloud environment. So, we have concluded our model perform faster in cloud environment than standalone environment although there are more resources available, which make computations possible in less time. Here, the comparative analysis visualized in Figure 11.

**TABLE 9.** Classification accuracy of different methods with existing work.

| Existing Methods | PIMA Dataset |
|---|---|
| | ACC (%) |
| LR [45] | 77.09 |
| NN [46] | 88.06 |
| XGBoost [47] | 88.00 |
| GA-stacking [48] | 88.88 |
| stacking [49] | 89.72 |
| KNN+ ERTC [50] | 85.06 |
| AWAIS [51] | 75.87 |
| NB [52] | 79.57 |
| DT [53] | 88.00 |
| Logistic Regression [60] | 78.00 |
| **PCA+ ELM (Proposed Model)** | **90.57** |

We have compared the classification accuracy of proposed model compared with existing models. We observed that our model performs best accuracy than other existing models. The comparison results are specified in Table 9.

## VI. CONCLUSION

In this proposed work, an Extreme Learning Machine (ELM) based automated diabetes prediction model with cloud-based environment has been suggested. The cloud computing technology provides continuous services at any time and from any location, so our proposed work provides the medical services in the rural work and detect the diabetes in the earlier stage. The extreme learning machine is using in this work because it avoids the local minima, fast convergence and simple as compare to other traditional classifiers. The proposed model has been simulated over standalone and cloud environment. The PCA technique has been used to reduce the feature dimension. The model provides better accuracy 90.57 % with 5 number of features.

The proposed eHealth system is employed as an "Application-as-a-Service" by using cloud computing techniques, providing different services like diagnosis detection, and assisting the pathology technician, doctors and decrease the mortality rate. By utilizing more resources in a cloud environment in the future, this approach can be developed further, potentially enhancing classification accuracy. Additionally, many ELM parameters can be modified to improve the effectiveness of the proposed framework. Moreover, the suggested model can be used in the field of image processing, where it can be used to a variety of tasks such character recognition, medical imaging, satellite images, and photo enhancement. In the future, we will apply some optimization technique and optimal feature selection methods to develop the generalization system performance and also provides cost-effective services to the real world.

## ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, Saudi Arabia for funding this research work through the project number "NBU-FFR-2023-0002".

## REFERENCES

[1] D. M. Muoio and C. B. Newgard, "Molecular and metabolic mechanisms of insulin resistance and $\beta$-cell failure in type 2 diabetes," *Nature Rev. Mol. Cell Biol.*, vol. 9, no. 3, pp. 193–205, Mar. 2008.

[2] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, "Classification and prediction of diabetes disease using machine learning paradigm," *Health Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–14, Dec. 2020.

[3] J. C. Prather, D. F. Lobach, L. K. Goodwin, J. W. Hales, M. L. Hage, and W. E. Hammond, "Medical data mining: Knowledge discovery in a clinical data warehouse," in *Proc. AMLA Annu. Fall Symp.* Bethesda, MD, USA: American Medical Informatics Association, 1997, p. 101.

[4] C. Azad and V. K. Jha, "Fuzzy min–max neural network and particle swarm optimization based intrusion detection system," *Microsyst. Technol.*, vol. 23, no. 4, pp. 907–918, Apr. 2017.

[5] C. Azad and V. K. Jha, "Data mining in intrusion detection: A comparative study of methods, types and data sets," *Int. J. Inf. Technol. Comput. Sci.*, vol. 5, no. 8, pp. 75–90, Jul. 2013.

[6] C. Azad and V. K. Jha, "A novel fuzzy min-max neural network and genetic algorithm-based intrusion detection system," in *Proc. 2nd Int. Conf. Comput. Commun. Technol.* New Delhi, India: Springer, 2016, pp. 429–439.

[7] C. Azad, S. Jain, and V. K. Jha, "Design and analysis of data mining based prediction model for Parkinson's disease," *Int. J. Comput. Sci. Eng.*, vol. 3, no. 3, pp. 181–189, 2014.

[8] M. A. Mohammed, K. H. Abdulkareem, S. A. Mostafa, M. K. A. Ghani, M. S. Maashi, B. Garcia-Zapirain, I. Oleagordia, H. Alhakami, and F. T. Al-Dhief, "Voice pathology detection and classification using convolutional neural network model," *Appl. Sci.*, vol. 10, no. 11, p. 3723, May 2020.

[9] A. Botta, W. De Donato, V. Persico, and A. Pescapé, "Integration of cloud computing and Internet of Things: A survey," *Future Gener. Comput. Syst.*, vol. 56, pp. 684–700, May 2016.

[10] R. Maskeliūnas, R. Damaševičius, and S. Segal, "A review of Internet of Things technologies for ambient assisted living environments," *Future Internet*, vol. 11, no. 12, p. 259, Dec. 2019.

[11] D. Połap and M. Woźniak, "Introduction to the model of the active assistance system for elder and disabled people," in *Proc. Int. Conf. Inf. Softw. Technol. (ICIST)*, in Communications in Computer and Information Science, vol. 639. Cham, Switzerland: Springer, 2016, Oct. 2016, pp. 392–403.

[12] B. Ayeni, O. Y. Sowunmi, S. Misra, R. Maskeliunas, R. Damaševicius, and R. Ahuja, "A web based system for the discovery of blood banks and donors in emergencies," in *Proc. Int. Conf. Intell. Syst. Design Appl. (ISDA)*, in Advances in Intelligent Systems and Computing, vol. 1181. Cham, Switzerland: Springer, Dec. 2019, pp. 592–600.

[13] B. D. Kanchan and M. M. Kishor, "Study of machine learning algorithms for special disease prediction using principal of component analysis," in *Proc. Int. Conf. Global Trends Signal Process., Inf. Comput. Commun. (ICGTSPICC)*, Dec. 2016, pp. 5–10.
[14] F. Esposito, D. Malerba, G. Semeraro, and J. Kay, "A comparative analysis of methods for pruning decision trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 476–491, May 1997.
[15] F. Liu and M. Brown, "Breast cancer recognition by support vector machine combined with Daubechies wavelet transform and principal component analysis," in *Proc. Int. Conf. ISMAC Comput. Vis. Bio-Eng. (ISMAC-CVB)*. Cham, Switzerland: Springer, 2019, pp. 1921–1930.
[16] F. Mohanty, S. Rup, B. Dash, B. Majhi, and M. N. S. Swamy, "Mammogram classification using contourlet features with forest optimization-based feature selection approach," *Multimedia Tools Appl.*, vol. 78, no. 10, pp. 12805–12834, May 2019.
[17] M. Dong, Z. Wang, C. Dong, X. Mu, and Y. Ma, "Classification of region of interest in mammograms using dual contourlet transform and improved KNN," *J. Sensors*, vol. 2017, pp. 1–15, Nov. 2017.
[18] I. Kavakiotis, O. Tsave, and A. Salifoglou, "Machine learning and data mining methods in diabetes research," *Comput. Struct. Biotechnol. J.*, vol. 15, no. 9, pp. 104–116, 2017.
[19] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," *Artif. Intell. Med.*, vol. 23, pp. 89–109, Aug. 2001.
[20] W. Cao, Z. Fang, G. Hou, M. Han, X. Xu, J. Dong, and J. Zheng, "The psychological impact of the COVID-19 epidemic on college students in China," *Psychiatry Res.*, vol. 287, May 2020, Art. no. 112934.
[21] Y. Sun, H. Wang, B. Xue, Y. Jin, G. G. Yen, and M. Zhang, "Surrogate-assisted evolutionary deep learning using an end-to-end random forest-based performance predictor," *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 350–364, Apr. 2020.
[22] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[23] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 879–892, Jul. 2006.
[24] D. Muduli, R. Dash, and B. Majhi, "Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach," *Biomed. Signal Process. Control*, vol. 59, May 2020, Art. no. 101912.
[25] K. Vembandasamy, R. Sasipriya, and E. Deepa, "Heart diseases detection using Naive Bayes algorithm," *Int. J. Innov. Sci., Eng. Technol.*, vol. 2, no. 9, pp. 441–444, 2015.
[26] S. Radhimeenakshi, "Classification and prediction of heart disease risk using data mining techniques of support vector machine and artificial neural network," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, New Delhi, India, Mar. 2016, pp. 3107–3111.
[27] A. Mohebbi, T. B. Aradottir, A. R. Johansen, H. Bengtsson, M. Fraccaro, and M. Mørup, "A deep learning approach to adherence detection for type 2 diabetics," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jeju, South Korea, 2017, pp. 2896–2899.
[28] T. Pham, T. Tran, D. Phung, and S. Venkatesh, "Predicting healthcare trajectories from medical records: A deep learning approach," *J. Biomed. Inform.*, vol. 69, pp. 218–229, May 2017.
[29] A. Askarzadeh and A. Rezazadeh, "Artificial neural network training using a new efficient optimization algorithm," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 1206–1213, Feb. 2013.
[30] N. M. Rao, K. Kannan, X.-Z. Gao, and D. S. Roy, "Novel classifiers for intelligent disease diagnosis with multi-objective parameter evolution," *Comput. Electr. Eng.*, vol. 67, pp. 483–496, Apr. 2018.
[31] C. Iwendi, S. A. Moqurrab, A. Anjum, S. Khan, S. Mohan, and G. Srivastava, "N-sanitization: A semantic privacy-preserving framework for unstructured medical datasets," *Comput. Commun.*, vol. 161, pp. 160–171, Sep. 2020.
[32] 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," *J. Healthcare Eng.*, vol. 2021, Oct. 2021, Art. no. 9930985.
[33] D. Das, D. R. Nayak, R. Dash, and B. Majhi, "An empirical evaluation of extreme learning machine: Application to handwritten character recognition," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 19495–19523, Jul. 2019.
[34] K. M. Ting, S. C. Tan, and F. T. Liu, "Mass: A new ranking measure for anomaly detection," *IEEE Trans. Knowl. Data Eng.*, pp. 1–13, 2009.
[35] S. Jangiti, E. S. Ram, and V. S. S. Sriram, "Aggregated rank in first-fit-decreasing for green cloud computing," in *Cognitive Informatics and Soft Computing*. Singapore: Springer, 2019, pp. 545–555.
[36] D. Muduli, R. Dash, and B. Majhi, "Automated diagnosis of breast cancer using multi-modal datasets: A deep convolution neural network based approach," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 102825.
[37] A. K. Dewangan and P. Agrawal, "Classification of diabetes mellitus using machine learning techniques," *Int. J. Engg. Appl. Sci.*, vol. 2, no. 5, pp. 145–148, May 2015.
[38] D. Muduli, R. Dash, and B. Majhi, "Fast discrete curvelet transform and modified PSO based improved evolutionary extreme learning machine for breast cancer detection," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 102919.
[39] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.
[40] M. A. Hall and G. Holmes, "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1437–1447, Nov./Dec. 2003.
[41] S. Jangiti and V. S. S. Sriram, "Scalable and direct vector bin-packing heuristic based on residual resource ratios for virtual machine placement in cloud data centers," *Comput. Electr. Eng.*, vol. 68, pp. 44–61, May 2018.
[42] L. Zhang and D. Zhang, "Domain adaptation extreme learning machines for drift compensation in E-nose systems," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 1790–1801, Jul. 2015.
[43] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: Algorithm, theory and applications," *Artif. Intell. Rev.*, vol. 44, no. 1, pp. 103–115, 2015.
[44] S. K. Sharma, A. Priyadarshi, S. K. Mohapatra, J. Pradhan, and P. K. Sarangi, "Comparative analysis of different classifiers using machine learning algorithm for diabetes mellitus," in *Proc. Int. Conf. Metaheuristics Softw. Eng. Appl.* Cham, Switzerland: Springer, 2022, pp. 32–42.
[45] A. Al-Zebari and A. Sengur, "Performance comparison of machine learning techniques on diabetes disease detection," in *Proc. 1st Int. Informat. Softw. Eng. Conf. (UBMYK)*, Nov. 2019, pp. 1–4.
[46] J. J. Khanam and S. Y. Foo, "A comparison of machine learning algorithms for diabetes prediction," *ICT Exp.*, vol. 7, no. 4, pp. 432–439, Dec. 2021.
[47] D. E. James and E. R. Vimina, "Machine learning-based early diabetes prediction," in *Intelligent Sustainable Systems*. Singapore: Springer, 2022, pp. 661–678.
[48] Y. Tan, H. Chen, J. Zhang, R. Tang, and P. Liu, "Early risk prediction of diabetes based on GA-stacking," *Appl. Sci.*, vol. 12, no. 2, p. 632, Jan. 2022.
[49] W. Wang, G. Xu, Y. Li, Z. Jin, L. Li, and Z. Dai, "Diabetes prediction model based on data enhancement and algorithm ensemble," *Proc. SPIE*, vol. 12168, Mar. 2022, Art. no. 121682Q.
[50] S. C. Gupta and N. Goel, "Enhancement of performance of $K$-nearest neighbors classifiers for the prediction of diabetes using feature selection method," in *Proc. IEEE 5th Int. Conf. Comput. Commun. Autom. (ICCCA)*, Oct. 2020, pp. 681–686.
[51] S. S. Sahan, K. Polat, H. Kodaz, and S. Günes, "The medical applications of attribute weighted artificial immune system (AWAIS): Diagnosis of heart and diabetes diseases," in *Proc. 4th Int. Conf. ICARIS*, Banff, AB, Canada, Aug. 2005, pp. 456–468.
[52] V. Chang et al., "Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms," *Neural Comput. Appl.*, pp. 1–17, 2022.
[53] M. O. Olusanya, R. E. Ogunsakin, M. Ghai, and M. A. Adeleke, "Accuracy of machine learning classification models for the prediction of type 2 diabetes mellitus: A systematic survey and meta-analysis approach," *Int. J. Environ. Res. Public Health*, vol. 19, no. 21, p. 14280, Nov. 2022.
[54] R. Spencer, F. Thabtah, N. Abdelhamid, and M. Thompson, "Exploring feature selection and classification methods for predicting heart disease," *Digit. Health*, vol. 6, Jan. 2020, Art. no. 205520762091477.
[55] M. Gabryel and R. Damaševičius, "The image classification with different types of image features," in *Proc. Int. Conf. Artif. Intell. Soft Comput.*, in Lecture Notes in Computer Science, 2017, pp. 497–506.
[56] D. Muduli, R. Dash, and B. Majhi, "Enhancement of deep learning in image classification performance using VGG16 with swish activation function for breast cancer detection," in *Proc. Int. Conf. Comput. Vis. Image Process.* Singapore: Springer, 2021, pp. 191–199.

[57] E. K. Bodur and D. D. Atsa'am, "Filter variable selection algorithm using risk ratios for dimensionality reduction of healthcare data for classification," *Processes*, vol. 7, no. 4, p. 222, Apr. 2019.
[58] M. Alwateer, A. M. Almars, K. N. Areed, M. A. Elhosseini, A. Y. Haikal, and M. Badawy, "Ambient healthcare approach with hybrid whale optimization algorithm and Naïve Bayes classifier," *Sensors*, vol. 21, no. 13, p. 4579, Jul. 2021.
[59] C. C. Olisah, L. Smith, and M. Smith, "Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective," *Comput. Methods Programs Biomed.*, vol. 220, Jun. 2022, Art. no. 106773.
[60] P. Rajendra and S. Latifi, "Prediction of diabetes using logistic regression and ensemble techniques," *Comput. Methods Programs Biomed. Update*, vol. 1, Jan. 2021, Art. no. 100032.

**SANTOSH KUMAR SHARMA** (Graduate Student Member, IEEE) received the M.Sc. degree in information technology from Panjab Technical University, in 2011, and the M.Tech. degree in computer science and engineering from NMIET, BPUT, Rourkela, Odisha, India, in 2013. He is currently pursuing the Ph.D. degree in computer science and engineering with C. V. Raman Global University, Bhubaneswar, Odisha. He is also a full-time Research Scholar with C. V. Raman Global University. He has more than ten years of teaching experience with NMET, Bhubaneswar, from 2011 to 2021, and Biju Patnaik Technical University, Odisha. He has two Indian patents and published two conference papers on artificial intelligence. His research interests include machine learning, cloud computing, deep learning, human–computer interaction, and clinical data analysis.

**ABU TAHA ZAMANI** is a Lecturer with the Department of Computer Science, Northern Border University, Saudi Arabia. He has published more than 17 papers in reputed journals and conferences. His current research interests include cloud computing, service selection, and optimization. He served as a reviewer for many reputed journals and conferences.

**AHMED ABDELSALAM** is an Assistant Professor of information systems and the Chairperson of the Board of Directors of the Saudi Computer Society. His key quality attributes are: Computational and Sustainable Higher Education Strategy Implementation in Saudi Arabia, Context-Aware Trust and Reputation Model for Fog-Based IoT Using Fuzzy TOPSIS, Multicriteria Decision Making for DevOps Data Quality Assessment Challenges Identification and Prioritization of Agile Requirements Change Management Success Factors in the domain of Global Software Development–A Systematic Study to Improve the Requirements Engineering Process in the domain of Global Software Development. His current research interests include ERP, CRM, knowledge management, e-government, e-business, business process reengineering, business intelligence, artificial intelligence (AI), the Internet of Things (IOT), digital transformation, machine learning, image processing, voice recognition, health informatics, cloud computing, and social networks applications.

**DEBENDRA MUDULI** received the M.Tech. degree in computer science and engineering and the Ph.D. degree from the National Institute of Technology, Rourkela, Odisha, India, in 2016 and 2022, respectively. He is an Assistant Professor with the Department of CSE, C. V. Raman Global University, Bhubaneswar, India. He has published many articles in reputed journals and conferences. His current research interests include cloud computing, pattern recognition, and medical image processing. He was a reviewer of many reputed journals and conferences.

**AMERAH A. ALABRAH** received the M.S. degree in computer science from Colorado State University, in 2008, and the Ph.D. degree in computer science from the College of Computer Science and Engineering, University of Central Florida, in 2014. Her research is mainly focused on computer and network security and more specifically on optimizing security measures for social media networks. She is currently an Assistant Professor with the College of Computer and Information Sciences, King Saud University, and a member of the Saudi Telecom Company Artificial Intelligence Research Fund.

**NIKHAT PARVEEN** received the B.Sc. degree in computer science and the M.C.A degree from Andhra University, Andhra Pradesh, India, in 2000 and 2003, respectively, and the Ph.D. degree from the Department of Computer Application, Integral University, Lucknow, Uttar Pradesh, India. She is currently an Associate Professor with the Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India. She has more than 12 years of teaching experience and six years of research experience. She has six national patents. Her research and publication interests include artificial intelligence, machine learning, security software, security testing, software engineering, and requirement engineering. She is also working in the area of soft computing, image analysis, big data analytics, and the IoT. Her research has been chronicled in over 30 journal publications and international conferences. She is a Life Time Member of CSI, ACM, IAENG, and IACSIT.

**SULTAN M. ALANAZI** is an Assistant Professor and Head of the Department of Computer Science at Northern Border University, Saudi Arabia. He holds a Ph.D. degree in computer science and the master's degree in IT from The University of Nottingham, UK, where he also worked as a Teaching Assistant with extensive experience in information technology. He has published several research papers in esteemed international journals and conferences. His research interests include AI, machine learning, natural language processing, cybersecurity, and recommender systems.