IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 13 May 2023, accepted 8 June 2023, date of publication 12 June 2023, date of current version 22 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3285247

RESEARCH ARTICLE

A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning

ABDULWAHAB ALI ALMAZROI[1], EMAN A. ALDHAHRI[2], SABA BASHIR[3], AND SUFYAN ASHFAQ[3]

[1]Department of Information Technology, College of Computing and Information Technology at Khulais, University of Jeddah, Jeddah 23445, Saudi Arabia
[2]Department of Computer Science and Artificial Intelligence, College of Computer Sciences and Engineering, University of Jeddah, Jeddah 23445, Saudi Arabia
[3]Department of Computer Science, Federal Urdu University of Arts, Sciences and Technology, Islamabad 44000, Pakistan

Corresponding author: Abdulwahab Ali Almazroi (aalmazroi@uj.edu.sa)

The authors extend their application to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through project number MoE-IF-UJ-22-04100409-X.

ABSTRACT Unfortunately, heart disease is currently the primary cause of mortality worldwide and its incidence is increasing. Detecting heart disease in its initial stages before a cardiac event takes place poses challenges. Huge amount of heart disease data is available in the health care sector such as in clinics, hospitals etc. However, this data is not intelligently handled to identify the hidden patterns. Machine learning techniques help in turning this medical data into useful knowledge. Machine learning is used to design such decision support systems (DSS) that can learn and improve from their past experiences. Recently, deep learning has gained the interest of industry and academics. The fundamental objective of this research activity is the precise diagnosis of heart illness. The suggested approach makes use of a Keras-based deep learning model to compute results with a dense neural network. The proposed model undergoes testing with various configurations of hidden layers in the dense neural network, ranging from 3 layers to 9 layers. Each hidden layer employs 100 neurons and utilizes the Relu activation function. To carry out the analysis, several heart disease datasets are utilized as benchmarks. The assessment encompasses both individual and ensemble models, and is performed on all heart disease datasets. Furthermore, using important measures like sensitivity, specificity, accuracy, and f-measure, the dense neural network is assessed across all datasets. The performance of different layer combinations varies across datasets due to varying attribute categories. Through extensive experimentation, the results of the proposed framework are analyzed. The study's conclusions show that, when applied to all heart disease datasets, the deep learning model suggested in this research paper achieves superior accuracy, sensitivity, and specificity compared to individual models and alternative ensemble approaches.

INDEX TERMS Machine learning, decision support system, deep learning, ensemble classifiers, heart disease diagnosis, accuracy, performance, cross validation.

I. INTRODUCTION

Heart disease, also referred to as cardiovascular illness, has become the primary cause of death worldwide and is becoming increasingly common. There are numerous symptoms associated with heart disease such as chest pain, sweating and fatigue, whereas mostly people feel nothing until a heart attack occurs [1]. According to the ranking performed by World Health Organization, Pakistan stands at 63rd level all around the world, where the ratio of heart disease is 110.65 per 10, 0000 from the deaths [2], [3].

To diagnose heart disease in patients, medical professionals use various techniques such as performing physical examinations, analyzing the patient's medical history, and conducting several medical tests. Normally a suspected heart disease patient is evaluated by using the clinical history, chest x-ray and physical examination, in spite of the fact that often some signs and isolated symptoms do not match up with these objective methods. Most of the people having heart attacks and heart strokes have not been identified as "at risk" by the medical experts and medical specialist. About

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

1/3rd of the patients have been misdiagnosed by medical specialists [1].

There are many symptoms of heart disease common to many other disease such as fatigue and chest pain which makes it difficult to diagnose at early stages. In addition, detecting heart disease has become a difficult task for many healthcare providers. This is because many patients do not show any symptoms until they suffer a sudden cardiac arrest [4], [5].

There are several disease management approaches such as nursing based intervention [6] and technological intervention [7] that have been evaluated by medical experts for boosting the health status of patients having heart disease. However, the complex and expensive implementation of these approaches do not make them a huge success.

Cardiac Biomarkers test has been emerged as cost effective methods to facilitate the medical professional for heart disease diagnosis. Discovering the biomarkers in the blood of heart disease patients has been preferred by many medical professionals. Cardiac biomarkers include hormones, enzymes and proteins. The biomarkers appear in the blood when the patient's heart is not getting enough oxygen. It shows that the person is getting a cardiac arrest.

However, the biomarkers are ambiguous in certain instances and occasionally additional tests are necessary to accurately diagnose heart disease. However, this risk assessment model is not good enough because it was built upon assuming linear relationship between each risk factor and heart disease outcome. Sometimes, the risk assessment of heart disease also depends on several other factors like smoking, diabetes, cholesterol and hypertension. Therefore, tremendous expertise is required by health care professionals to analyze the patient. Many times over complexity of the patient may lead the doctors to make errors in order to identify the heart disease [8].

Huge amount of medical data is available in the health care sector such as clinics, hospitals etc. However, this data is not intelligently handled to extract useful knowledge. Based on a survey of 37 million patient records, it is estimated that 195,000 people in the US died due to the unavoidable medical errors. If sickness diagnosis is carried out using effective machine-learning approaches, medical errors may be greatly minimized [9].

Machine learning can be employed to extract valuable insights from vast amounts of clinical data in order to diagnose heart disease. These advanced computational methods prove highly efficient in distinguishing between healthy individuals and those affected by heart disease, establishing a clear demarcation. Machine learning-based intelligent algorithms are used by Clinical Decision Support Systems (CDSS) to aid in the diagnosis of cardiac disease. It's essential to note that CDSS relies on foundational knowledge of the data to perform precise analysis. A decision support system provides the necessary knowledge for the prediction of any disease [13] and accuracy level to diagnosing the heart disease can be improved in health care community by using such machine learning based decision support systems.

Supervised and unsupervised machine learning are two subcategories of machine learning techniques. The goal of supervised learning is to learn about a function by using a dataset that contains categorized input and matching intended output values. The learning algorithms are used for the classification of unlabeled data by using the classified data set as a basis to predict the disease. One use for supervised machine learning is the diagnosis of illnesses. Numerous techniques, including Naïve Bayes, Linear Regression, Decision Trees, and Neural Networks, among others [11], may be used to do this. However, unsupervised learning does not contain labeled outputs so its goal is to draw inferences from data sets containing the input values. The learning algorithm is not been provided with any labels. There are some pattern reorganization problems in which training data includes a set of input values without any corresponding output values. In such supervised learning problems the objective is may be to find out groups of similar examples inside the data which is known as clustering [14].

In the realm of CDSS, there exist two distinct classifications that shape its very essence. The categorizations of CDSS, in a similar fashion to the paths diverging in a digital forest, lead to two branches: CDSS that are knowledge-based and non-knowledge-based. Among these, the based on knowledge CDSS stands out as a beacon of accuracy. It was carefully crafted to help users traverse the immense sea of data and information, ensuring that accurate judgments are made at the appropriate moment and in the appropriate context. Like a skilled conductor orchestrating a symphony of information, it harmonizes the elements of proper data utilization and timely wisdom, painting a portrait of accuracy in the realm of medical diagnosis. This approach basically focuses on the diagnosis of several diseases [12]. On the other hand non-knowledge based CDSS only covers a narrow list of symptoms such as diagnosis of a single disease. Therefore, powerful and reliable Knowledge based CDSS are used to predict the disease especially heart disease by saving the diagnosis time and increasing the prediction accuracy [9].

In conventional machine learning, a set of hand-crafted relevant features are analyzed by algorithm whereas, in deep learning the algorithm is provided with the raw data and it decides on its own that what features are relevant. Deep learning bypasses the need of features as input from any user. The performance of deep learning networks will normally improve as the amount of data provided for training has been increased. Deep learning algorithm learns from the experience just like the human brain and each time performs the task with a better outcome. The majority of the current decision-making systems rely on machine learning, requiring human intervention to input manually designed characteristics. However, deep learning-based systems excel as the forefront of technology and are employed as clinical systems to support disease diagnosis. Moreover, mostly existing systems are evaluated on single datasets which can show biasness towards that dataset whereas proposed deep learning based system is evaluated on 4 different heart disease datasets obtained from different online repositories which show that proposed system have consistent high performance and it can be applied on any medical dataset [15].

The technology has been lifted by some quick advancement in fast data storage, parallelization and computational power. Additionally, it has a strong predictive capability and semantic understanding of the input data. It has been applied successfully in dimensionality reduction [15], motion modeling [16], and object detection etc. The algorithms of deep learning use deep architectures or multiple layer architectures to remove inherent features of data having multiple levels of abstraction. It focuses on learning feature hierarchy methods and they can be used to find large amount of structures in the data [17].

In heart disease prediction, Khan, M. A. [18] presented a model that employed Internet of Things (IoTs) and an altered deep convolutional neural network (CNN) for the task. Sensor data is used for analysis and proposed method performed better than existing methods. The working of proposed framework may be extended using fully wearable devices. In a study [19], the author introduced a framework for the Internet of Medical Things aimed at diagnosing heart disease. The results showed significant improvements in accuracy, precision, and recall compared to existing models. However, improvements can be made using further optimization techniques. Khan, M. A. et al. [20] incorporated secure methods to introduce a secure framework. The efficacy of the model is validated through the performance of a formal security analysis. Authentication and encryption are two important factors of security.

Creating a deep learning model that can detect heart disease is the primary objectives. The effectiveness of the suggested approach is evaluated using four already-existing datasets on cardiac disease. Deep neural networks with numerous layers are used to obtain great analytical accuracy. Data imputation techniques are employed to make refine form. Two different comparisons are used to show the suggested model's effectiveness: one with individual classifiers and another with ensemble classifiers.

In order to make the rest of the work easier to grasp, it is separated into the following sections: A thorough literature assessment of the most recent deep learning frameworks for heart disease is presented in Section II. The suggested technique is fully explained in Section III. Section IV deals with the presentation of results and experimental assessment. Section V wraps up the work and considers potential lines of inquiry for future work.

II. LITERATURE REVIEW

This section presents the critical analysis of research papers focused on machine learning and deep learning based clinical decision support systems for heart disease diagnosis.

In [21] Chae et al. used deep learning and big data to predict infectious diseases. The study focuses on predicting scarlet fever, malaria and chickenpox by comparing the performance of Long Short-Term Memory (LSTM) and Deep Neural Network (DNN) with the Autoregressive Integrated Moving Average (ARIMA). Dataset of 576 observations is collected from KCDC, Naver Data Lab, KMA and Twitter. The data consist of four types which includes social media data, query data, and temperature. Deep neural network is a technique which uses a back-propagation algorithm and is used to predict time series data. The prediction of infectious diseases is performed using ARIMA where the analysis of non- stationary time series data is performed and proved that it can be implemented on any time series data. It displays the changes of fluctuated data in detail.

In [22] Raju et al. focused on automatic prediction of diabetes using deep learning algorithm. The authors proposed a framework to classify the stage of diabetic retinopathy using the funduscopic images. A convolution approach based on deep neural network is used for the network training on benchmark heart disease data set. The dataset was archived from Kaggle data repository. The classification analysis indicates the high values of sensitivity and specificity for different stages of diabetic retinopathy. Training of network is performed using 35,000 images approximately which achieved a sensitivity and specificity of 80.28% and 90.29% respectively. Approximately 8,810 images were used to train the model and it has achieved an accuracy of 93.28%. It was observed that in order to improve the classification accuracy, many other factors can also be considered such as family history, demographics and past history etc. Furthermore, images with poor quality need to be filtered out to reduce the error rate.

Reddy and Khare [23] designed a hybrid heart disease system using Oppositional Firefly with BAT (OFBAT) and Ruled Based Fuzzy Logic (RBFL) for heart disease prediction. Locality Preserving Projection (LRP) is used for the selection of desired features. It helps to design classification model by using fuzzy logic system. Subsequently, the rules are defined for fuzzy system form the sample data. After that, OFBAT algorithm is used for the selection of relevant and important rules. Afterwards, the fuzzy rules and membership functions help to design a fuzzy system where classification can be performed. Ultimately, the testing is performed on the publicly accessible UCI data sets (Cleveland, Hungarian and Switzerland). Their results display that RBFL achieved an accuracy of 78% for UCI data sets. The proposed methodology can be improved using deep learning.

In [24] Khateeb et al. proposed a framework for heart disease diagnosis. A number of machine learning techniques are evaluated on publicly available UCI data set (303 records). The authors divided their work into six cases. In 1st case, the classifiers are tested without using feature reduction. In 2nd case, feature reduction is applied by using only 7 relevant attributes out of 14. In 3rd case, the accuracies are calculated by removing most generic features like sex, age and resting blood pressure. In 4th case, re sampling is performed on data sets by using Weka tool and the accuracies are calculated for seven essential attributes. In 5th case, re sampling is applied to all the 14 attributes. Ultimately, In 6th case, the accuracies are calculated by applying Synthetic Minority Over Sampling Technique (SMOTE) in Weka tool. The KNN approach achieved best results of approximately 80% accuracy. The proposed approach is limited to the use of single dataset.

In [25] Verma et al. proposed a framework for heart disease diagnosis along with its important factors identification. The risk factors are identified using different methods as Particle Swam Optimization (PSO) and K-means clustering method. After successfully identifying the risk factors, supervised learning algorithms are used for the classification. Three medical datasets are used for the experimentation. These datasets are Cleveland heart disease data, India patient's data and 335 instances collected from Indhira Gandhi Medical Collage. The analysis of results indicates that Multinomial Logistic Regression performs better on all the datasets as compared to other classifiers.

In [26] Chen et al. presented a decision support system for the medical professionals. The proposed model utilized the available data in order to diagnose the heart disease. The proposed framework is based on Learning Vector Quantization (LVQ) algorithm for the prediction and analysis. Firstly, 13 significant attributes are selected from the data set. Secondly, heart disease classification is performed using Artificial Neural Network on these selected attributes. Ultimately, a heart disease user friendly system is developed. The analysis of results is indicated using Receiver Operating Characteristics (ROC) curve whereas 80% accuracy was achieved. If the textual data is available then text mining can be used to further improve the performance of proposed model as it can perform prediction of unstructured data which can be used for heart disease diagnosis.

Jabbar et al. [27] utilized Naïve Bayes classifier in order to improve the predictive accuracy for early diagnosis of heart disease. Irrelevant and redundant attributes are removed by using discretization and genetic search. Feature selection is performed using Genetic Algorithm (GA) where least ranked attributes are removed from the dataset. Eventually, the performance of Naïve Bayes and other approaches are compared with each other. It has achieved 86.29% accuracy for heart disease statlog data set.

In [28] Srinivas et al. also used machine learning techniques and introduced a clinical decision support system. The proposed methodology is based on three basic classification algorithms. Benchmark datasets are used for the experimentation and analysis. Data preprocessing is also performed by the extended version of Naïve Bayes. Missing values imputation is also performed to improve the quality of data where mean and mode methods are used. The proposed framework is advantageous over other frameworks as it can easily answer the complex queries for heart disease diagnosis.

Chitra and Sinivasagam [29] proposed a model by using the supervised learning algorithms for heart disease diagnosis using the medical records of patients. The authors used Cascaded Neural Network (CNN) to classify the information exists in patient's medical records. The classification of heart disease is performed by the 13 selected attributes as an input to CNN classifier in the classification stage. The proposed technique performance is evaluated on the records of 270 patients to test the efficiency of classifier. The efficiency of proposed CNN is clearly analyzed by comparing the results with state of the art methods. This system can provide help to physicians in the diagnosis of heart disease.

In [30] Helmy et al. presented a machine learning based decision support system for heart disease diagnosis. First, individual classifiers are used for the training. Then, Bagging algorithm is used to ensemble the results of individual classifiers. The analysis of results indicates the heterogeneous classifiers fusion produced more effective results as compared to other ensembles and single classifiers. However, the performance of proposed approach may be analyzed using many other benchmark and real time datasets.

Table 1 shows the summary of state of art techniques for heart disease diagnosis.

There are number of machine learning algorithms used for heart disease diagnosis. The accuracy of diagnosis can be further improved using deep learning framework.

III. PROPOSED METHODOLOGY

The proposed methodology is divided into two stages:

1- Data acquisition and preprocessing
2- Classification Models and Proposed Model

In data acquisition, the heart disease datasets are collected from a benchmark online available machine learning repository. The preprocessing includes replacement of missing values with actual values, normalization, standardization, outlier's detection and elimination and then finally removing duplicates. The missing values in data must be handled carefully as the performance of machine learning and deep learning algorithms may deteriorate by having missing values in the data. If an attribute of a dataset contains missing values then these values are replaced by taking the mean of that specific attribute. If a certain feature of a dataset contains more than 50 percent of missing or NAN values then it should be discarded. The datasets are organized and converted into a uniform format using the normalization and standardization. After that the outliers in the datasets have been detected and replaced by the binning method where attribute data is divided into bins and then mean is used to replace the bin values. Finally, the dataset containing duplicates has been cleaned by removing the duplicate values. If a dataset contains record of the same patient several times then that duplicate considered only one time.

The second phase includes the computation of results on the acquired and preprocessed datasets by applying the

TABLE 1. An overview of state of the art techniques for heart diseases prediction.

Reference	Year	Technique	Advantages	Disadvantages
Hassan, D. et al [38]	2023	Logistic Regression	Improved performance utilizing deep neural network and PCA	Ensemble algorithm may be employed for high classification
Dileep, P. et al [39]	2023	Bi directional LSTM	Good for identifying risk factors of heart disease	Optimization and other feature selection techniques may be used
Khan, M. A. [18]	2020	Modified Deep CNN	IoT based framework for heart disease prediction with high accuracy	Evaluated on single heart disease dataset, fully wearable devices may be used
Khan, M. A. et al [19]	2020	MSSO and ANFIS	Utilized optimization technique to improve the performance	Limited to use of other feature selection and optimization techniques
Singh, A. et al [31]	2020	Support Vector Machine, Decision Tree, Linear Regression, K Nearest Neighbor	Analyzed KNN is better as compared to other single classifiers	Single machine learning classifiers are used, evaluation is limited to one dataset
Maji, S. et al [32]	2019	Decision Tree, Artificial Neural Network	Proposed a hybridized technique	Evaluation is performed on single dataset
Chauhan, A. et al [33]	2018	Evolutionary Learning	Rule based method is introduced	Evaluation is limited to single dataset, Statistical evaluation is not performed
Reddy, G. T. et al [23]	2017	Radial Basis Network Link Network	Optimized the rules of fuzzy logic system	Achieved accuracy below 80%
Khateeb, N. at al [24]	2017	K Nearest Neighbor	proposed a model for heart disease diagnosis	Limited to use of single machine learning classifier and one dataset
Verma, L. et al [25]	2016	FURI, Multi layer Perceptron, Multiple Linear Regression, C4.5	Hybrid model is proposed, Evaluation is performed on benchmark and non-invasive clinical data	Evaluated on single benchmark heart disease data, more clinical data instances may be required
Sharma, P. et al [34]	2016	Rule based Classifier	Rule based system to predict risk level patients	Limited to use of single classifier
Bashir, S. et al [10]	2015	Ensemble Bagging	Proposed multi level classifiers framework	other ensemble techniques may be incorporated
Bashir, S. et al [35]	2014	Ensemble Vote	Majority vote based ensemble classifier	Single dataset is used for evaluation

individual classifiers such as Nearest Neighbors, Gaussian Process, Linear SVM, Decision Tree, Naive Bayes, QDA, AdaBoost, Bagging, Boosting and proposed Dense Neural Network. 10-Fold Cross validation is applied on the datasets before feeding them to the classification models. The method is used to randomly divide the datasets into 10 chunks. The 9 of them are used for the training of models and 1 of them is used for the testing purpose. The process is repeated 10 times each time a different 1/10th part is reserved for the testing of classification models. The Scikit-Learn based classification models are used for the computation of results for individual classifiers. The Scikit-Learn is basically a python library. The performances of individual classifiers are measured on the basis of sensitivity, accuracy and specificity. Following classification algorithms are used for the prediction of heart disease.

A. K-NEAREST NEIGHBOR

K- Nearest Neighbor is a classification approach that determines how a data point is classified based on its distance from a closest group of points [3]. The distance between data points is calculated by Euclidean distance formula which is given as follows:

$$d\,(p,q) = \sqrt{\sum\nolimits_{i=1}^{n} (qi - pi)\,2} \tag{1}$$

where p and q represents the points in the n-space. The data point is classified on the basis of minimum value of

Euclidean distance which is calculated between a data point from different closest group of points.

B. SUPPORT VECTOR MACHINE

It is also a classification algorithm for identifying the classes in given data. The algorithm is trained with data that is already classified into two classes. The class identification of new data point is done using this built model. It is not only used to classify the objects but also get the margins between them on plot as wide as possible [4]. Following support vectors are used for the method:

$$W_0^T x + b_0 = 1 or W_0^T x + b_0 = -1 \tag{2}$$

C. DECISION TREE

Decision tree is basically used for pattern reorganization and cleaning of data. There are number of features in a given dataset and perhaps a part of them are less important as compared to others. Gini Index can be used to eliminate the irrelevant and weakly relevant features. In order to construct decision tree, each node is selected by using the calculation of conditional probabilities The features having low Gini Index is selected and using those selected features rules are generated. The Decision tree can be used in bioinformatics such as for different disease prediction and diagnosis [5]. The following formula is used for the calculation of Gini Index.

$$G = \sum_{I=1}^{C} p(i) * (1 - p(i)) \tag{3}$$

whereas c represents the classes number and p(i) represents the probability of each class.

D. Naïve BAYES

The naïve bayes classifier presumes that the existence of a specific feature in a class is independent to the existence of any other feature. The class feature is the only requirement for this classifier as the other class features are unrelated to each other [6]. The given formula is used for the calculation of this classifier.

$$P(C_K|X) = \frac{P(C_K) * P(X|C_K)}{P(X)} \tag{4}$$

where CK is the probability of a class per the given feature X.

E. GAUSSIAN PROCESS

It is basically a random process which means a group of variables indexed by space or time. The process is a generalized form of Gaussian probability distribution. The functions in Gaussian probability distribution used to sum up the random variables whereas in Gaussian process the properties of a function are summarized. It is useful in statistical modeling and used to take benefits from the properties that are inherited from the normal distribution [7]. The given function is used to calculate the gaussian process.

$$\mathrm{f}(\mathrm{x}) = \mathrm{a.exp}((\mathrm{x}-\mathrm{y})^2/2\mathrm{c}^2) \tag{5}$$

where a represents the curve's peak height, y represents the center of peak and c indicates the standard deviation.

F. QUADRATIC DISCRIMINATE ANALYSIS (QDA)

The classifier is used for the non-linear separation of data. The classifier uses a bayes rule to generate the quadratic decision based boundary and to fit the data to the class conditional densities. Basically QDA is not that much different from Linear Discriminate Analysis (LDA) expect that in QDA all classes may not share the identical covariance matrix. The data fitting mechanism of QDA is better as compared to LDA because it allows more flexibility in terms of covariance matrix [8]. The following formula is used to calculate QDA.

$$\mathrm{f_K(X)} = \frac{1}{2}\log\left|\sum \mathrm{K}\right| - (\mathrm{X}-\mathrm{U_K})^{\mathrm{T}} \sum\nolimits_{\mathrm{K}}^{-1} (\mathrm{X}-\mathrm{U_K}) + \log \pi_{\mathrm{K}} \tag{6}$$

where $\sum$ K is covariance matrix, which is calculated for each class k=1, 2, 3…k.

G. BOOSTING

It is a basically a group of machine learning algorithms where a strong model is generated by combining different weaker models. The model is usually built using the decision trees. The model is famous because of its effective ability for the classification of datasets. It is primarily a Adaboost technique merged with the weighted minimization. After the minimization of weights the weighted inputs and classifiers are recalculated. The main object of boosting is to minimize the loss which is basically the variation between the actual value and the predictive value [10]. The following equation is used to calculate boosting.

$$\mathrm{F_b}(\mathrm{x_i}) = \mathrm{F_{b-1}}(\mathrm{x_i}) - \gamma - \times \nabla \mathrm{L}(\mathrm{y_i}, \mathrm{F}(\mathrm{x_i})) \tag{7}$$

where $\mathrm{F_b}$ indicates classifier version, γ is learning rate, ∇ represents the gradient, $\mathrm{x_i}$ represents the number of observations and L indicates the loss function between the actual and predicted value.

H. BAGGING

It is a kind of ensemble approach in machine learning (also known as Bootstrap Aggregation) used to improve the model performance by combining the outputs from many learner. It is often used for the removal of variance within a noisy data. In the approach, random subsets of training data are selected. After generating the several data samples, train them independently on various machine learning models. Finally, overall prediction is generated by combining their predictions together. It is mainly used in machine learning for classification tasks such as Decision Trees and Naïve Bayes etc. It may also be applied to regression algorithms but it has been found to be more effective for classification tasks [11]. It is calculated using the given equation.

$$\mathrm{f_{bag}} = \mathrm{f_1}(\mathrm{x}) + \mathrm{f_2}(\mathrm{x}) + \mathrm{f_3}(\mathrm{x}) + \ldots\ldots\ldots.\mathrm{f_n}(\mathrm{x}) \tag{8}$$

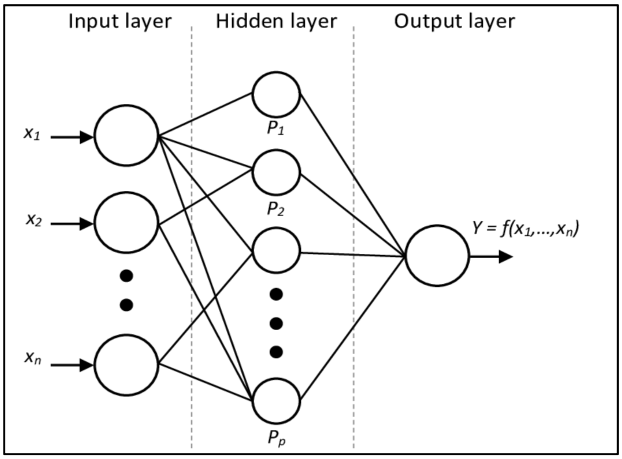

FIGURE 1. **Architecture of dense neural network.**

The terms on the right hand side represents individual learners and the term on the left hand side indicates the bagging prediction.

I. DENSE NEURAL NETWORK

A Dense Neural Network (DNN) is basically an Artificial Neural Network (ANN) contains one or more hidden layers however there is only one input layer and one output layer. The DNN consists of several components such as Neurons, Biases, Weights and Functions. These components work on the principle of human brain and used to train the machine learning algorithms [12]. The neurons inspired by the biological neurons of the human brain and activate to perform specific action just like the action performed by the human body in response to a certain condition. Summarily, each neuron receives a dot product of random weights and multiple inputs from the input layer which is then added with a constant bias. The output value is passed to an activation function and it checks whether the network is accurately predict the particular pattern or not. If the prediction is not accurate the weights would be adjusted by the algorithm [13]. The following generalized function is used to calculate it.

$$F_i = f(\sum_{i=1}^{n} x_i w_i + b) \tag{9}$$

where x_i indicates the number of input features, w_i indicates their corresponding weights, b stands for bias and f indicates the activation function that is applied to weighted sum of input features.

Figure 1 shows an architectural representation of a feed-forward dense neural network. There are three layers in the network and each layer has its own functionality. The hidden layers are basically used to reduce the error rate; therefore its number can vary depending on the performance.

J. WORKING OF PROPOSED DEEP LEARNING MODEL

The keras based deep learning model is used for the computation of results for dense neural network. Keras is basically an API of deep learning running at the top of platform like TensorFlow. Keras provides high-level API while TensorFlow provides both high and low level API. The following dense neural network model is defined with five layers. The first layer is input layer which is used to define the shape of input features. The second, third and fourth are hidden layers. The last layer is output layer.

```
input = Input(shape=(None,13))
d1 = Dense(units=100, activation='relu')(input)
d1 = Dense(units=100, activation='relu')(d1)
d1 = Dense(units=100, activation='relu')(d1)
output = Dense(units=1, activation='sigmoid')(d1)
model = tf.keras.models.Model(inputs=input, outputs=out
put)
```

The proposed model is tested with different combination of hidden layers, started from 3 to 9. Each hidden layer uses 100 neurons and activation function as relu. The function is used to convert the inputs to outputs and it allows the model to perform better and learn faster. Finally, the output layer having 1 unit of neuron and activation function as sigmoid as the classification problem is binary. The last line of code is Keras functional API that converts this whole architecture into a model.

```
model.compile(loss='binary_crossentropy', metrics=['accura
cy', tf.keras.metrics.Recall(name="Sensitivity"),tf.keras.met
rics.SpecificityAtSensitivity(0.5,name="Specificity"),
tfa.metrics.F1Score(num_classes=1, threshold=0.5) ], optim
izer=Adam(learning_rate=0.00001))
```

For the compilation of the proposed model, some addition parameters are required for better evaluation of model and finding the best set of weights. The binary cross entropy used as a loss function because the problem is a binary classification. The objective is to minimize the loss and adjust the model weights to train the model better. The accuracy, sensitivity and specificity metrics are used to measure the performance of proposed model. The Adam optimizer is used which is basically a famous version of gradient descent. It is used to optimize the loss and gives the best results in most of the problems. The 0.00001 learning rate is selected on the basis of experiments performed with different learning rates like 0.1, 0.01, 0.001, 0.0001 and 0.00001.

```
model.fit(x=X_train, y=y_train, batch_size=2, epochs=120)
```

After the compilation, the model is trained with batch size = 2 and epochs = 120. The batch size and epochs are found to be best on the basis of experiments performed by varying the epochs and batch sizes. The proposed framework is shown in Figure 2.

The algorithm of proposed methodology is given in table 2.

IV. EXPERIMENTAL EVALUATION AND MEASURES

The proposed framework is evaluated using different performance metrics to show the effectiveness of proposed work. Each performance metric is given below [36] and [37].

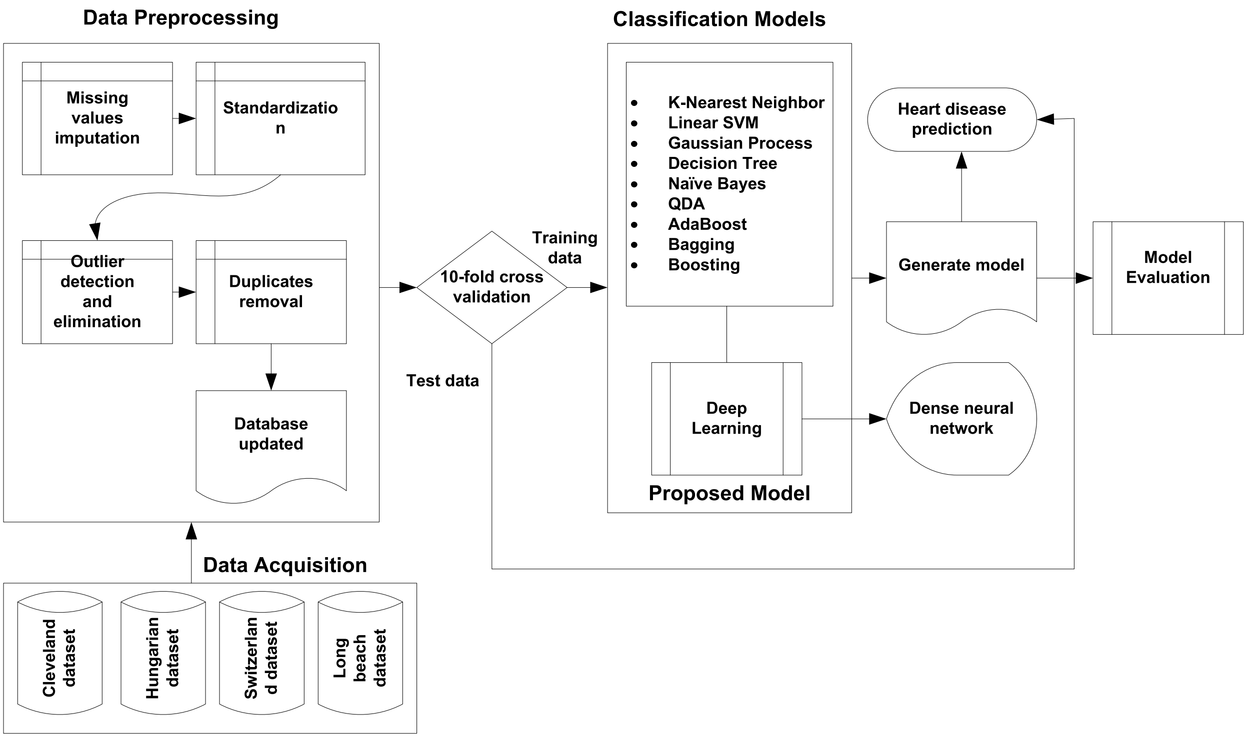

FIGURE 2. Architectural representation of proposed methodology.

TABLE 2. Algorithm of proposed methodology.

Algorithm: A Novel Heart Disease Prediction Algorithm
Input: Heart Disease Dataset
Results: Disease Prediction (Yes/No)
    for each instance in the dataset
      Perform Data Pre-Processing
        Missing values imputation (Instance)
        Standardization (Instance)
        Outlier detection (Instance)
        Duplicates handling (Instance)
      Return refined data

    for each refined data
      Perform 10-fold cross validation on refined data
        if Training data then
          Train classification algorithms on training data
          Train Deep learning algorithm on training data
            Return Trained model
        else
        Apply Trained model on Test data
        Calculate Evaluation measures
        Return Statistical Results

A. ACCURACY

Accuracy is the ratio of correct assessments made by a classification model as compared to the total assessments which is executed on the test data [9].

$$Accuracy = \frac{Truenegatives + Truepositives}{positives + negatives} \tag{10}$$

B. SENSITIVITY

The percentage of correct assessments labeled as correct by a classification model. It is calculated using the following formula [9].

$$Sensitivity = \frac{Truepositives}{Falsepositives + Truepositives} \tag{11}$$

C. SPECIFICITY

The percentage of wrong assessments labeled as wrong by a classification model. It is calculated using the following equation [9].

$$Specificity = \frac{Truenegatives}{Falsenegatives + Truenegatives} \tag{12}$$

D. F-MEASURE

F-Measure is the ratio between sensitivity and specificity.

$$F - Measure = \frac{2 * Sensitivity * Specificity}{Sensitivity + Specificity} \tag{13}$$

V. ANALYSIS OF RESULTS

The results are computed by executing the code on Google Colaboratory. The google colab basically allows the user to execute python code using the browser. The colab platform provides 12GB of RAM. The proposed approach starts off with evaluating the results for individual classifiers. The individual classifiers which performed best are selected and tested on benchmark heart disease datasets.

The heart disease datasets are gathered from the online available machine learning UCI data repository [1] which is approved by many different researchers. Four benchmark heart disease datasets are used namely Cleveland, Hungarian, Long Beach and Switzerland. Each dataset have specific features that will be used to determine if the patient is sick with heart disease or not. The label attribute (num) which is basically the output feature consist of categorical values in the form of healthy and sick. The healthy values have been replaced by 1 and sick values have been replaced by 0.

The detail of each dataset is mentioned below.

A. CLEVELAND DATASET

The Cleveland dataset has a total of 14 features. The 13 features are input features and they are independent of each other. The last column is an output feature which is basically a label attribute and it depends on the input features like if it is 0 it means patient is sick with heart disease otherwise not. The dataset contains 303 instances.

B. HUNGARIAN DATASET

The Hungarian dataset contains 12 features and 294 instances. The 13 input features are independent of each other while the output feature named "num" depends on input features. The label attribute "num" contains number values (0-4). If the values are 0 it means the patient is sick with heart disease otherwise not.

C. SWITZERLAND DATASET

The Switzerland dataset contains 13 input features and 1 output feature which is basically a dependent feature. The dataset contains 123 instances and the goal attribute contains values in the form 0 to 4. The values 0 indicate the existence of heart disease while the other values indicate the absence of heart disease.

D. LONG-BEACH DATASET

The long-beach dataset contains 13 independent input features and 1 output feature which basically depends on input feature. The goal feature contains 200 instances and it values range from 0 to 4. If the values are 0 it means the patient is sick with heart disease otherwise not.

10-fold cross validation is performed to split each dataset into training set and test set. In each fold, one subset is reserved for testing and remaining 9 subsets are used for training. Average accuracy is calculated to show the final result. This is the accuracy of that trained classifier. This process is repeated for each dataset and for each classifier.

Table 3 and Table 4 show the description of heart disease datasets. All datasets contain same number of features however the difference between them is range of values in each attribute, attribute characteristics and output attribute range.

Table 5 presents results of individual classifiers on the benchmark heart disease datasets. Each classifier is evaluated against different performance metrics to show the effectiveness. Moreover, execution time is also calculated to determine the time taken by the classifier in training and testing.

TABLE 3. Datasets features and description.

Features Name	Description
Age	Age stated in years
Sex	Gender type
CP	Type of chest pain
Trestbps	Blood pressure level at resting state
Chol	Total cholesterol in blood (mg/dl)
Fbs	Level of fasting blood sugar > 120 mg/dl
Restecg	Resting electrocardiographic results
Thalach	Max heart rate achieved
Exang	Exercise induced angina
Oldpeak	ST depression induced by exercise relative to rest
Slope	Slope of the peak exercise ST segment
Ca	Number of major vessels colored by flourosopy
Thal	6= fixed defect; 3=Normal; 7= reversible defect

The comparison of results indicates that mostly ensemble classifiers take longer execution time when compared with individual classifiers. However, performance of ensemble classifiers is better.

Table 6 shows results of deep learning algorithm on all datasets using multiple layers ranging from 3 to 9 layers. Dense neural network is used for the evaluation. The hidden layers combination is selected based on the best performing combination for benchmark heart disease datasets.

For Cleveland dataset, Table 6 shows that 4-layer combination of hidden layers has significant good performance as compared to other combination of hidden layers. It has statically better accuracy and smaller execution time as compared to other combination of hidden layers.

For Hungarian dataset, Table 6 indicates that 3-layer combination of hidden layers has statistically significant results as compare to other combination of hidden layers. Statistic shows that it has good accuracy, specificity, sensitivity and F-score compared to other combination of hidden layers. Moreover, it has smaller execution time for the implementation of network.

For Switzerland dataset, 8-layer combination of hidden layers performed significantly better in terms of specificity, accuracy, sensitivity, f-score and execution time when its statistics are compared with other combination of hidden layers.

For Long Beach dataset, 9-layer combination of hidden layers has statistically good result in comparison with other combinations. The statistics show that it has performed specificity, accuracy, sensitivity, f-score and execution time better based on accuracy, specificity, sensitivity, f-score and execution time.

The analysis of results indicates that different combination of layers performs well for different datasets due to different range of values. Moreover, it is also analyzed that deep learning performs better as compared to individual and ensemble classifiers. The reason behind better results is that automatic feature extraction and model training without human intervention.

TABLE 4. Comparison of heart disease datasets.

Cleveland Dataset			
Data Variation	Multivariate	**Count of samples**	303
Type of Attribute	Categorical, Integer, Real	**Number of Attributes**	14

Switzerland Dataset			
Data Variation	Multivariate	**Count of samples**	123
Type of Attribute	Categorical, Integer, Real	**Number of Attributes**	14

Hungarian Dataset			
Data Variation	Multivariate	**Count of samples**	294
Type of Attribute	Integer, decimal	**Number of Attributes**	12

Long Beach Dataset			
Data Variation	Multivariate	**Count of samples**	200
Type of Attribute	Categorical, Integer, Real	**Number of Attributes**	14

TABLE 5. Evaluation of individual classifiers results.

	Cleveland Dataset					Hungarian Dataset				
Techniques	Acc	Spec	Sens	F-Score	Exe Time	Acc	Spec	Sens	F-Score	Exe Time
Nearest Neighbors	80.54	80	80.58	80.31	0.0028	76.89	62.63	73.8	73.43	0.002
Linear SVM	79.83	75.65	79.55	79.53	0.0003	81.31	66.27	78.04	78.46	0.0002
Gaussian Process	79.5	74.94	79.2	79.06	0.0005	82.01	66.45	78.66	78.74	0.0004
Decision Tree	72.59	71.97	72.53	72.44	0.0002	75.8	75.54	75.68	74.04	0.0001
Naive Bayes	81.23	79.28	81.16	81.01	0.0002	82.31	76.72	81.12	80.36	0.0001
QDA	46.84	90	50	31.87	0.0016	41.56	80	50	28.96	0.0015
AdaBoost	74.52	60.54	73.54	73.07	0.0048	81.29	66.18	78	78.69	0.0042
Bagging	79.51	77.08	79.38	79.16	0.0041	81.65	67.27	78.54	78.82	0.0039
Boosting	77.19	57.63	75.74	75.37	0.0005	63.95	0	50	39	0.0003

	Switzerland Dataset					Long Beach Dataset				
Nearest Neighbors	62.56	57.33	62.23	61.63	0.0014	61.5	67.27	61.08	60.78	0.002
Linear SVM	54.48	0	50	35.23	0.0002	53.49	1	50	34.83	0.0002
Gaussian Process	61.66	44.66	60.3	59.2	0.0003	63.5	65.9	63.12	62.34	0.0004
Decision Tree	58.65	59	59.61	57.44	0.0001	61	66.54	60.93	59.03	0.0001
Naive Bayes	63.58	78.33	64.64	62.03	0.0002	50	8.36	53.12	39.29	0.0002
QDA	45.51	1	50	31.23	0.001	53.49	1	50	34.83	0.0012
AdaBoost	59.23	38.66	57.66	53.1	0.0039	62.5	83.09	60.93	59.12	0.0052
Bagging	61.79	46.66	60.71	59.59	0.0025	58.5	65.18	58.03	57.08	0.0028
Boosting	55.32	13.99	51.99	40.91	0.0004	62	85.81	60.07	57.71	0.0004

Figure 3 depicts the graphical representation of proposed framework comparison with individual classifiers for Cleveland dataset. Accuracy, sensitivity, specificity and f-measure of individual classifiers and ensemble classifiers are compared with deep learning. It is clear from the given figure that proposed deep leaning has performed significantly better in terms of accuracy, specificity, sensitivity and f-score.

TABLE 6. Evaluation of deep learning results.

	Cleveland Dataset					Hungarian Dataset				
Deep Learning	Acc	Spec	Sens	F-Score	Exe Time	Acc	Spec	Sens	F-Score	Exe Time
3-layer	78.86	76.42	81.02	80.59	0.58	83.03	69.27	90.9	87.37	0.5
4-layer	82.5	78.57	85.95	84.16	0.51	81.66	68.36	89.29	86.21	0.51
5-layer	79.54	76.42	82.35	81.29	0.57	80.66	68.36	87.74	85.34	0.55
6-layer	81.51	76.37	85.95	83.43	0.5	82.02	68.36	89.85	86.59	0.55
7-layer	81.18	77.14	84.77	82.92	0.49	83.01	69.18	90.93	87.32	0.49
8-layer	80.82	77.85	83.49	82.43	0.52	80.32	69.27	86.66	84.9	0.49
9-layer	82.49	77.08	87.24	84.41	0.6	81.02	68.36	88.27	85.74	0.55

	Switzerland Dataset					Long Beach Dataset				
3-layer	55.25	12.66	90.95	68.42	0.5	63.5	44.27	55.11	57.51	0.51
4-layer	50.77	13.66	91.19	69.18	0.5	62	44.36	53.66	55.83	0.54
5-layer	59.42	17.66	93.8	71.26	0.51	63	52.63	56	56.98	0.54
6-layer	63.97	34.66	87.85	72.69	0.55	61	47.72	53.77	54.83	0.55
7-layer	61.85	35	77.61	68.13	0.52	63	38.18	57.33	58.54	0.52
8-layer	64.29	45	80.47	70.69	0.54	63.99	45.72	63.66	61.62	0.53
9-layer	60.06	42.66	74.52	66.49	0.58	64	49.72	62.44	61.23	0.56

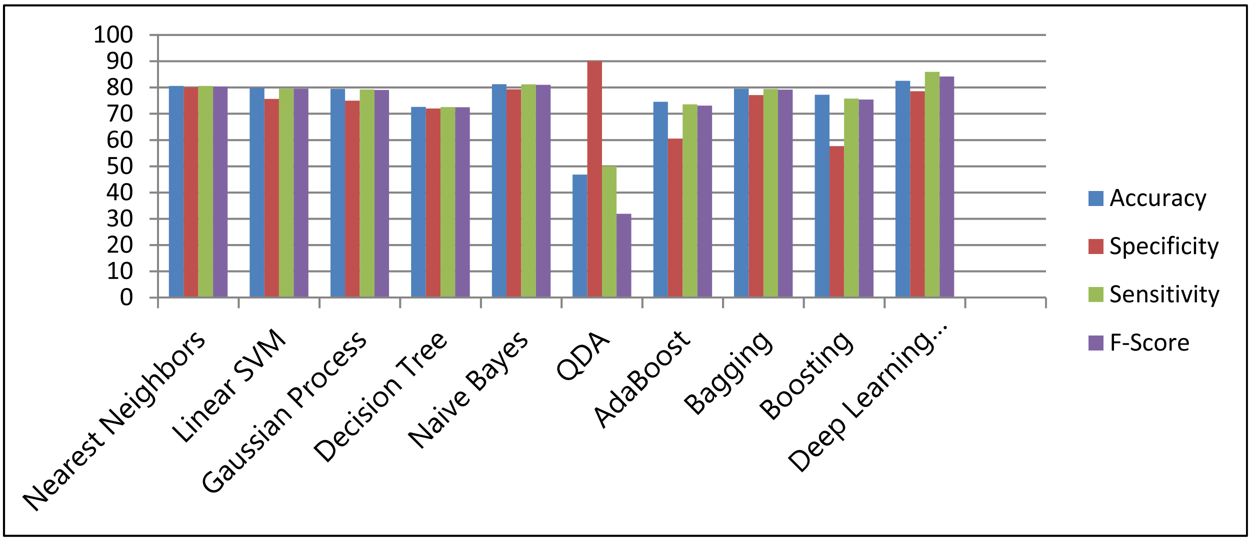

FIGURE 3. Graphical comparison of deep learning with individual and ensemble classifiers for Cleveland dataset.

Figure 4 shows the graphical comparison of proposed framework deep learning with individual and ensemble classifiers. The comparison shows that deep learning has performed better as compared to other classifiers for Hungarian dataset when comparison is performed using Accuracy, sensitivity, specificity and f-measure.

For Switzerland dataset, Figure 5 indicates that proposed deep learning has achieved better prediction accuracy when compared with other individual and ensemble classifiers. The statistic shows that results are extremely significant in comparison with other techniques.

The given Figure 6 shows that that deep learning has better results for Long Beach dataset in comparison with other classifiers. The proposed technique can predict heart disease with better accuracy, sensitivity, specificity and f-measure as compared to individual and ensemble classifiers.

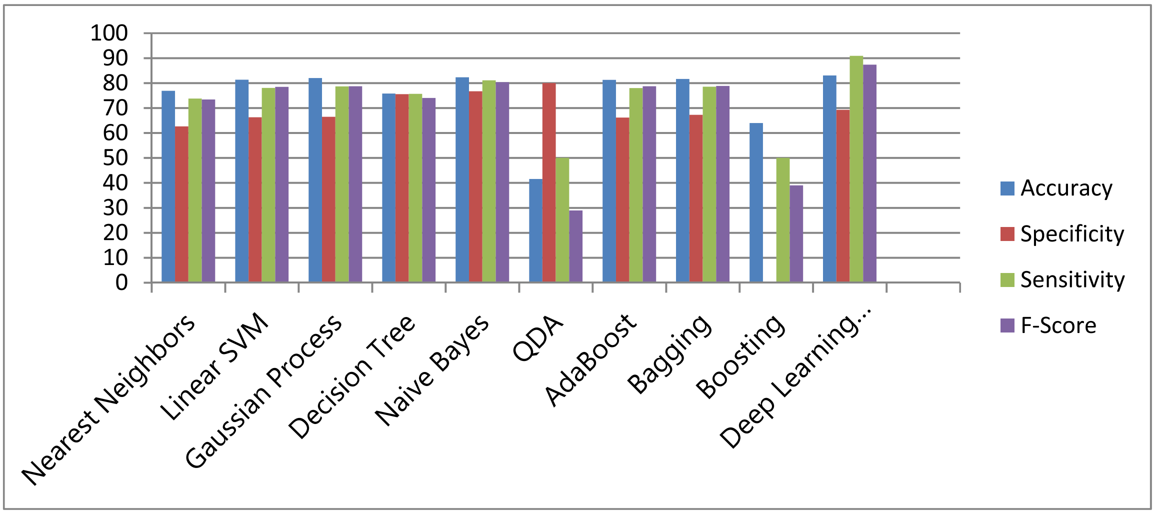

FIGURE 4. Graphical comparison of deep learning with individual and ensemble classifiers for Hungarian dataset.

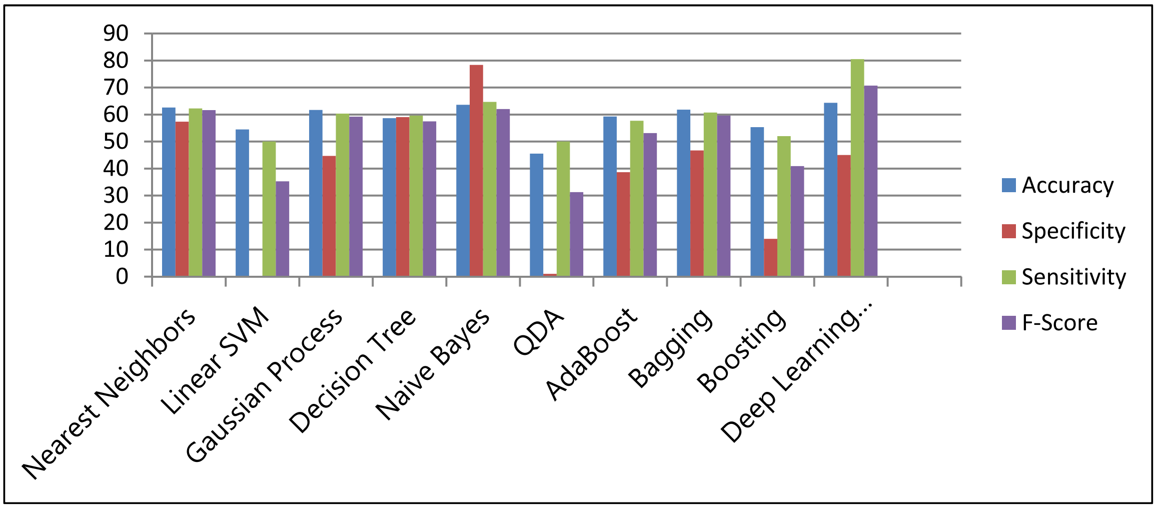

FIGURE 5. Graphical comparison of deep learning with individual and ensemble classifiers for Switzerland dataset.

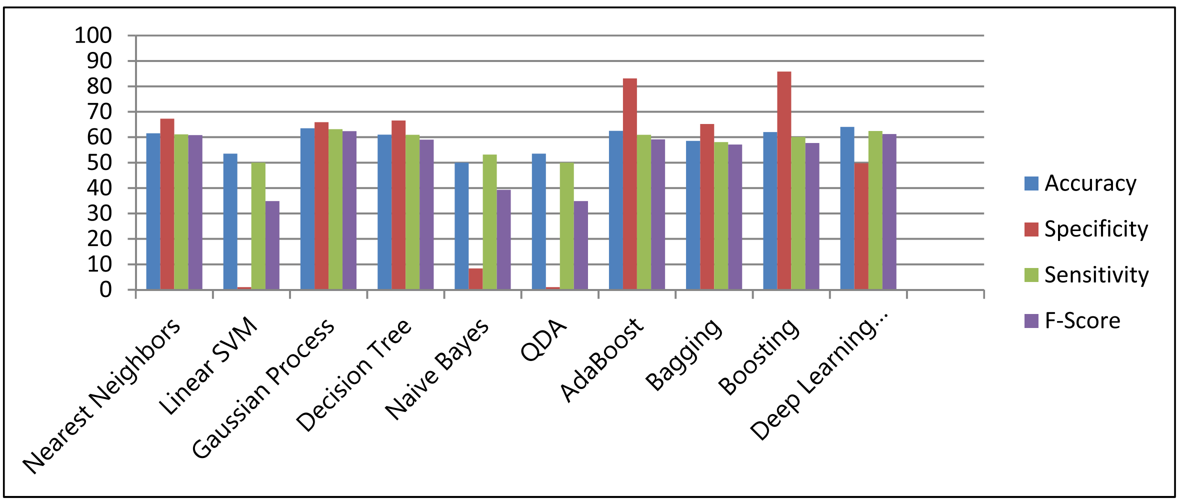

FIGURE 6. Graphical comparison of deep learning with individual and ensemble classifiers for Long Beach dataset.

State of the Art comparison of proposed framework is also performed with other techniques and results are shown in Table 7. Most of the existing frameworks are applied on single or 2 heart disease datasets. The proposed framework is evaluated to four different heart disease datasets having different range/type of values and features.

TABLE 7. State of the art comparison with proposed technique.

Reference	Techniques	Accuracy
Singh, A. et al [45]	Support Vector Machine, Decision Tree, Linear Regression	83%, 79%, 78%
Maji, S. et al [43]	Decision Tree, Artificial Neural Network	78%, 77%
Chauhan, A. et al [44]	Evolutionary Learning	58%
Reddy, G. T. et al [34]	Radial Basis Network Link Network	78%
Khateeb, N. at al [10]	K Nearest Neighbor	80%
Proposed	Deep Learning	83%

The proposed framework showed consistent high performance in all datasets which indicates that proposed deep learning framework can be applied on any dataset for heart disease prediction. It is not biased towards any dataset or any range of values. In all datasets deep learning showed more than 80% accuracy which indicates an acceptable level of performance for medical decision support system. This high performance is due to hidden layers of deep learning model where error rate is reduced. The higher accuracy of proposed deep learning framework indicates that it can predict heart disease efficiently.

VI. CONCLUSION AND FUTURE WORK

Deep learning is a mean for effective and accurate heart disease diagnosis and prediction. The proposed model performed significantly better in terms of accuracy, sensitivity and specificity as compared to other techniques. In the future, we would like to strengthen this approach by using the images data of heart disease patients. The images data will be collected using the laboratory examinations and imaging. Moreover, Convolution Neural Network (CNN) will be applied on the images data to diagnose the heart disease with maximum accuracy. The major advantage of using the CNN on given images data is that it detects the most important features automatically. Furthermore, some additional performance metrics will be used for the evaluation of model such as confusion matrix, PR curve and ROC curve.

Additionally, the CNN model can be tested on combined structured and unstructured data. It may boost the performance and accuracy power of CNN algorithms for heart disease prediction.

ACKNOWLEDGMENT

The authors extend their application to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through project number MoE-IF-UJ-22-04100409-X.

REFERENCES

[1] N.-S. Tomov and S. Tomov, "On deep neural networks for detecting heart disease," 2018, *arXiv:1808.07168*.
[2] World Heath Organization. (Jun. 7, 2014). *Cardio Vascular Diseases (CVDs)*. [Online]. Available: http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
[3] Worldlifeexpectancy. (2014). *World Health Raniknngs Live Longer Live Better*. [Online]. Available: http://www.worldlifeexpectancy.com/cause-of-death/coronary-heart-disease/by-country/
[4] C. Sowmiya and P. Sumitra, "Analytical study of heart disease diagnosis using classification techniques," in *Proc. IEEE Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS)*, Mar. 2017, pp. 1–5.
[5] A. Kumar, P. Kumar, A. Srivastava, V. D. A. Kumar, K. Vengatesan, and A. Singhal, "Comparative analysis of data mining techniques to predict heart disease for diabetic patients," in *Proc. Int. Conf. Adv. Comput. Data Sci.* Singapore: Springer, 2020, pp. 507–518.
[6] Y.-J. Huang, M. Parry, Y. Zeng, Y. Luo, J. Yang, and G.-P. He, "Examination of a nurse-led community-based education and coaching intervention for coronary heart disease high-risk individuals in China," *Asian Nursing Res.*, vol. 11, no. 3, pp. 187–193, Sep. 2017.
[7] R. Hasan, "Comparative analysis of machine learning algorithms for heart disease prediction," in *Proc. ITM Web Conf.*, vol. 40, 2021, p. 03007.
[8] S. Khan and S. T. Rasool, "Current use of cardiac biomarkers in various heart conditions," *Endocrine, Metabolic Immune Disorders-Drug Targets*, vol. 21, no. 6, pp. 980–993, Jun. 2021.
[9] S. Bashir, A. A. Almazroi, S. Ashfaq, A. A. Almazroi, and F. H. Khan, "A knowledge-based clinical decision support system utilizing an intelligent ensemble voting scheme for improved cardiovascular disease prediction," *IEEE Access*, vol. 9, pp. 130805–130822, 2021.
[10] S. Bashir, U. Qamar, and F. H. Khan, "BagMOOV: A novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 305–323, Jun. 2015.
[11] V. Vives-Boix, D. Ruiz-Fernández, A. de Ramón-Fernández, D. Marcos-Jorquera, and V. Gilart-Iglesias, "A knowledge-based clinical decision support system for monitoring chronic patients," in *Proc. Int. Work-Conf. Interplay Between Natural Artif. Comput.* Cham, Switzerland: Springer, 2017, pp. 435–443.
[12] S. Bashir, Z. S. Khan, F. Hassan Khan, A. Anjum, and K. Bashir, "Improving heart disease prediction using feature selection approaches," in *Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2019, pp. 619–623.
[13] A. J. Aljaaf, D. Al-Jumeily, A. J. Hussain, P. Fergus, M. Al-Jumaily, and K. Abdel-Aziz, "Toward an optimal use of artificial intelligence techniques within a clinical decision support system," in *Proc. Sci. Inf. Conf. (SAI)*, Jul. 2015, pp. 548–554.
[14] K. Chitra and D. Maheswari, "A comparative study of various clustering algorithms in data mining," *Int. J. Comput. Sci. Mobile Comput.*, vol. 6, no. 8, pp. 109–115, 2017.

[15] Y. Wang, H. Yao, and S. Zhao, ''Auto-encoder based dimensionality reduction,'' *Neurocomputing*, vol. 184, pp. 232–242, Apr. 2016.

[16] M. P. T. Kaandorp, S. Barbieri, R. Klaassen, H. W. M. Laarhoven, H. Crezee, P. T. While, A. J. Nederveen, and O. J. Gurney-Champion, ''Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients,'' *Magn. Reson. Med.*, vol. 86, no. 4, pp. 2250–2265, Oct. 2021.

[17] *The Enterprisers Project*. Accessed: Dec. 25, 2022. [Online]. Available: https://enterprisersproject.com/article/2019/7/deep-learning-explained-plain

[18] M. A. Khan, ''An IoT framework for heart disease prediction based on MDCNN classifier,'' *IEEE Access*, vol. 8, pp. 34717–34727, 2020.

[19] M. A. Khan and F. Algarni, ''A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS,'' *IEEE Access*, vol. 8, pp. 122259–122269, 2020.

[20] M. A. Khan, M. T. Quasim, N. S. Alghamdi, and M. Y. Khan, ''A secure framework for authentication and encryption using improved ECC for IoT-based medical sensor data,'' *IEEE Access*, vol. 8, pp. 52018–52027, 2020.

[21] S. Chae, S. Kwon, and D. Lee, ''Predicting infectious disease using deep learning and big data,'' *Int. J. Environ. Res. Public Health*, vol. 15, no. 8, p. 1596, Jul. 2018.

[22] M. Raju, V. Pagidimarri, R. Barreto, A. Kadam, V. Kasivajjala, and A. Aswath, ''Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy,'' in *Proc. MEDINFO*, 2017, pp. 559–563.

[23] G. T. Reddy and N. Khare, ''An efficient system for heart disease prediction using hybrid OFBAT with rule-based fuzzy logic model,'' *J. Circuits, Syst. Comput.*, vol. 26, no. 4, Apr. 2017, Art. no. 1750061.

[24] N. Khateeb and M. Usman, ''Efficient heart disease prediction system using K-nearest neighbor classification technique,'' in *Proc. Int. Conf. Big Data Internet Thing*, Dec. 2017, pp. 21–26.

[25] L. Verma, S. Srivastava, and P. C. Negi, ''A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data,'' *J. Med. Syst.*, vol. 40, no. 7, p. 178, Jul. 2016.

[26] A. H. Chen, S. Y. Huang, P. S. Hong, C. H. Cheng, and E. J. Lin, ''HDPS: Heart disease prediction system,'' in *Proc. Comput. Cardiol.*, Sep. 2011, pp. 557–560.

[27] M. A. Jabbar and S. Samreen, ''Heart disease prediction system based on hidden naïve Bayes classifier,'' in *Proc. Int. Conf. Circuits, Controls, Commun. Comput. (I4C)*, Oct. 2016, pp. 1–5.

[28] K. Srinivas, B. K. Rani, and A. Govrdhan, ''Applications of data mining techniques in healthcare and prediction of heart attacks,'' *Int. J. Data Mining Techn. Appl.*, vol. 7, no. 1, pp. 172–176, Mar. 2018.

[29] R. Chitra and V. Seenivasagam, ''Heart attack prediction system using cascaded neural network,'' in *Proc. Int. Conf. Appl. Math. Theor. Comput. Sci.*, 2013, p. 223.

[30] T. Helmy, S. M. Rahman, M. I. Hossain, and A. Abdelraheem, ''Non-linear heterogeneous ensemble model for permeability prediction of oil reservoirs,'' *Arabian J. Sci. Eng.*, vol. 38, no. 6, pp. 1379–1395, Jun. 2013.

[31] A. Singh and R. Kumar, ''Heart disease prediction using machine learning algorithms,'' in *Proc. Int. Conf. Electr. Electron. Eng.*, Feb. 2020, pp. 452–457.

[32] S. Maji and S. Arora, ''Decision tree algorithms for prediction of heart disease,'' in *Information and Communication Technology for Competitive Strategies*. Singapore: Springer, 2019, pp. 447–454.

[33] A. Chauhan, A. Jain, P. Sharma, and V. Deep, ''Heart disease prediction using evolutionary rule learning,'' in *Proc. 4th Int. Conf. Comput. Intell. Commun. Technol. (CICT)*, Feb. 2018, pp. 1–4.

[34] P. Sharma, K. Saxena, and R. Sharma, ''Heart disease prediction system evaluation using C4.5 rules and partial tree,'' in *Computational Intelligence in Data Mining*, vol. 2. New Delhi, India: Springer, 2016, pp. 285–294.

[35] S. Bashir, U. Qamar, and M. Y. Javed, ''An ensemble based decision support framework for intelligent heart disease diagnosis,'' in *Proc. Int. Conf. Inf. Soc.*, Nov. 2014, pp. 259–264.

[36] A. A. Almazroi, H. Alamin, R. Sujatha, and N. Z. Jhanjhi, ''A web-based model to predict a neurological disorder using ANN,'' *Healthcare*, vol. 10, no. 8, p. 1474, Aug. 2022.

[37] M. H. Alkinani, A. A. Almazroi, M. Adhikari, and V. G. Menon, ''Design and analysis of logistic agent-based swarm-neural network for intelligent transportation system,'' *Alexandria Eng. J.*, vol. 61, no. 10, pp. 8325–8334, Oct. 2022.

[38] D. Hassan, H. I. Hussein, and M. M. Hassan, ''Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis,'' *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104019.

[39] P. Dileep, K. N. Rao, P. Bodapati, S. Gokuruboyina, R. Peddi, A. Grover, and A. Sheetal, ''An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm,'' *Neural Comput. Appl.*, vol. 35, no. 10, pp. 7253–7266, Apr. 2023.

ABDULWAHAB ALI ALMAZROI received the M.Sc. degree in computer science from the University of Science, Malaysia, and the Ph.D. degree in computer science from Flinders University, Australia. He is currently an Associate Professor with the Department of Information Technology, College of Computing and Information Technology at Khulais, University of Jeddah, Saudi Arabia. His research interests include parallel computing, cloud computing, wireless communication, and data mining.

EMAN A. ALDHAHRI received the B.Sc. degree in computer science from Taibah University, Madinah, Saudi Arabia, in 2006, the M.Sc. degree in computer science from Southern Illinois University, Carbondale, IL, USA, in 2014, and the Ph.D. degree in computer science from The University of Memphis, Memphis, TN, USA, in 2019. He is currently an Assistant Professor with the Department of Computer Science and Artificial Intelligence and the Vice Dean of Research and Innovation with the College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia.

SABA BASHIR received the Ph.D. degree in software engineering from the College of Electrical and Mechanical Engineering (CEME), National University of Science and Technology, Pakistan. She is currently an Associate Professor with the Department of Computer Science, Federal Urdu University of Arts, Sciences and Technology, Islamabad, Pakistan. She is a Certified Engineer. She has published many research articles in top-tier peer-reviewed international journals in the field of data mining/machine learning. Her research interests include predictive systems, health informatics, artificial intelligence, and machine learning.

SUFYAN ASHFAQ received the B.S. degree in computer science from the Federal Urdu University of Arts, Science and Technology (FUUAST), Islamabad, Pakistan, and the B.Sc. degree from Punjab University, Lahore. He is currently pursuing the M.S. degree in computer science with FUUAST. His research interests include machine learning and deep learning.