International Journal of Computer Sciences and Engineering
Vol.**12**, Issue.**1**, pp.**08-15**, January **2024**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Article

Heart Disease Prediction System Using Convolutional Neural Networks

V. Krishnaiah[1]

[1]Dept. of Computer Science and Engineering, Neil Gogte Institute of Technology, Affiliated by Osmania University**,** Kachawanisingaram (Village), Uppal, Hyderabad, Telangana State, India, 500088.

Email: drvkrishnaiah@ngit.ac.in

Received: 15/Nov/2023; **Accepted**: 22/Dec/2023; **Published**: 31/Jan/2024. **DOI**: https://doi.org/10.26438/ijcse/v12i1.815

Abstract: Now a days, based on different reasons heart diseases are increasing rapidly. If we find out or identify the heart diseases in human beings at an early stage, it is easy to prevent the disease and help the patients. Even though cardiologists and health centers gather relevant data and information every day, but, not applying the knowledge of machine learning algorithms to retrieve valuable of prediction. The main objective of this research is to predict and classify heart diseases by using proposed convolutional neural network classifier. In this classification of evaluation process, feed forward process and back propagation methods will be applied in between the hidden layers. Due to this, the proposed CNN classifier gives best accuracy. By applying this trained classifier has identified the given data, which are either normal or abnormal. So, the entire research has been implemented in Python which produced good results.

Keywords: Deep Learning, Classification, Convolutional Neural Network, Heart Disease.

1. Introduction

Classification is one of the basic approches in computer science and other research technologies. It can evaluate the bases on many other computer science technology tasks such as object identification, image segmentation of and object prediction. The task of categorizing object into one of several predefined classes known as object classification. By using machine learning techniques, data set can be classified, and falls under the category of deep learning. Deep learning is a type of neural network algorithms in which each layer is responsible for extracting one or more features of the data sets.

A neural network is a computational model that is similar to a human brain. It is a collection of nodes known as neurons. These nodes are organized into layers where each neuron in the layer takes some input processes and passes the output to the neuron in the next layer. Different layers may perform different kinds of transformations. Data transfers from the input layer (first layer) to the output layer (last layer) by traversing various hidden layers. The *Convolutional Neural Network (CNN)* classifier used for improving the accuracy of image data set for classification [1]. The bellow Fig: 1 show the basic structure of an neural network.

The neural network can generate efficient classification rules. *Convolution neural network* classifier is a multilayer perceptron that is the special design for identification of 2-dimensional data information. Always have more layers: input layer, convolution layer, sample layer, and output layer. Deep learning refers to the shining branch of machine learning that is based on learning levels of representations.

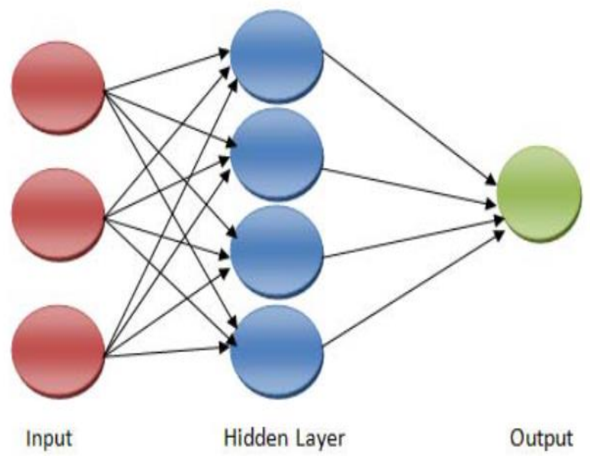

Figure 1: Neural Network

The *CNN* is one kind of deep neural network. To perform classification task of heart disease dataset, the neural network is trained using convolutions algorithm. The experiment is conducted with heart disease dataset by considering the single and multilayer neural network methods. The proposed algorithm gives detailed analysis of the process of *CNN* both the forward process and back propagation.

In this work, incorporated a *Convolutional Neural Network* based on object classifier which can identify and separate the heart disease data records of normal from that of abnormal. The CNN classifier performed well for performance measures of *Accuracy, Mean Value Error, Sensitivity, Specificity, F-Measure, False Positive Rate, False Negative Rate, and Overall Error Rat*e, and compared with existing data mining techniques for better accuracy.

The remainder of this article has been organized as follows: Related work has been briefly introduced in **Section 2**. The data source of heart disease data sets has been presented in **Section 3**. Experimental method/procedure/design of *CNN* and its architecture has been elaborated in **Section 4**. Results and discussion of *CNN* with existing classifiers using heart disease data sets mentioned in **Section 5**. Finally, concluding remarks.

2. Related Work

In [2], certain fusion methods have been built to diagnose CVDs sideways with its sternness. Machine learning (ML) techniques like artificial neural network, SVM, Logistic Regression, Decision Tree, Random Forest, and AdaBoost have been used to the heart disease dataset to predict diagnose. Random ended tester was executed as of the class inequity in multiclass classification. To increase the performance of classification, a weighted score fusion method was occupied. Initially, the methods were trained. After training, two techniques choice was joint applying a weighted sum rule. Some of these fusion methods have been established from the six ML techniques. The output was likely in the performance parameter. The proposed method has been investigated with dissimilar test training ratios for binary and multiclass classification problems, and for both of them, the fusion methods performed well. The maximum accuracy for multiclass classification was found as 75%.

In [3], projected k-nearest neighbor founded heart disease prediction method, and author calculated the performance of the proposed method. Further, the output has been calculated for predictive performance of the proposed method. Heart disease data sets are obtained from Kaggle machine learning data repository. The dataset contains of 1025 observations of which 499 or 48.68% is heart disease negative and 526 or 51.32% is heart disease positive. Lastly, the performance of KNN is examined on the test set. The output of performance examination on the experimental outputs on the Kaggle heart disease data repository displays that the accuracy of the KNN is 91.99%.

In [4], a correlative examination was established by applying ML techniques. Together cardiovascular and diabetes disease datasets were applied for classification. Dissimilar ML techniques similar XGBoost, Random Forest, and weighted ensemble methods applied to predict disease. The vital features that found maximum to the dataset were recognized. Finally, the performance parameters had been exposed to detect the outputs. Ensemble methods expressed marginally well accuracy than further methods. The academics also anticipated to use the method in a real-life situation to verify the risk of the disease happening. Author used ten-fold cross-validation, and its accuracy was marginally less.

Correlative examination of numerous diseases has been completed by applying Machine Learning Techniques [5, 6, 7, 8, 9]. Two data mining techniques have been applied [10] for the diagnosis of cardiovascular disease. Between them, the Support Vector Machine (SVM) exhibited a robust competence. The accuracy of SVM, and Artificial Neural Network was 74.42% and 73.64%, correspondingly [11, 12, 13, 14].

It [15], depends on the test bang values, diagnose a possible solutions. Over assessing the obtainable data gathered, and predicted whether the patient has heart disease or diabetes using the model. In this work, projected Naive Bayes technique applied for classifying the data set, and it reached accurate output.

In [16], proposed Machine Learning technique of SVM. This technique predicts the heart disease depends on the assumed features like sex, age, pulse rate etc. Here, SVM reached the greatest accuracy and reliable outputs then compared to the other techniques. And also evaluated performance measure with better accuracy, specificity, and sensitivity then compared to the other machine learning techniques.

3. Data Source

The data sets of Cleveland and Statlog are utilized. Cleveland data set comprises 303 patient records of heart disease data [17], and the Statlog data set comprises 270 patient records of heart disease data [18]. A total of 573 patient records are chosen for valuation of the proposed prediction scheme. The data from the patient records will possess 13 input attributes, namely, age, sex, cp, t-rest-bps, chol, Rest-ecg, fbs, thalach, exang, old-peak, solpe, Ca, thal. The detail descriptions of these attributes are provided in the given below Table 1.

Data Sets and Attributes

The data set was categorized with 3 attributes: Key, Predictable, and Input attributes as given below:

1. *Key Attribute:* Patient Id: Patient's Recognition Number.
2. *Predictable Attribute:* Diagnosis:
   Value 1 ≤ 50 % (no heart disease);
   Value 0 ≥ 50 % (has heart disease).
3. *Input Aattributes* (Heart Disease): Table 1.

Here, a key attribute of a patient record may be patient identification number (Patient ID), which is specified for every patient and can be easily distinguishable. The bellow 13 attributes of data of a patient are said to be as input attributes which may have some personal information. Along with these attributes there is one more attribute for every record and can be helpful to identify it even among a number of similar records also termed as key attributes. This database contains 76 attributes, but all published experiments refer to using a subset of 13 of them.

Table 1: Input Attributes for Prediction System

Variable Name	Attribute	Descriptions	Values
F1	Age	Age in years	Continuous
F2	Sex	Male or Female	1 = male; 0 = female
F3	Cp	Chest Pain Type	1 = typical type I 2 = typical type angina 3 = non-angina pain 4 = asymptomatic
F4	T rest bps	Resting blood pressure	Continuous value in mm hg.
F5	Chol	Serum cholesterol	Continuous value in mg/dl.
F6	Restecg	Resting electrographic results	0 = normal 1 = containing ST_T sign odd 2 = left ventricular hypertrophy
F7	Fbs	Fasting blood sugar	1 ≥ 120 mg/dl. 0 ≤ 120 mg/dl.
F8	Thalach	Maximum heart rate attained	Continuous value
F9	Chol	Exercise make angina	0 = no; 1 = yes
F10	Oldpeak	ST despair make by exercise virtual to rest	Continuous value
F11	Slop	Slope of the crest exercise ST fragment	1 = un sloping 2 = fat 3 = down sloping
F12	Ca	Number of main vessels tinted by fluoroscopy	0 - 3 value
F13	Thal	Defect type	3 = normal 6 = fixed 7 = reversible defect

In particular, the Cleveland, Statlog databases are used by Machine Learning researchers. The "goal" field refers to the presence of heart disease in the patient. It is integer value from 0 (no presence) to 4. Experiments with this database have concentrated on simply attempting to distinguish presence (values 1, 2, 3, 4) from absence (value 0). The names and social security numbers of the patients were recently removed from the database, replaced with dummy values.

4. Experimental Method/Procedure/Design

This *Convolutional Neural Network (CNN)* has input layers, output layers and hidden layers. The hidden layers consist of convolutional layer, flattened layer and a fully connected layer. The bellow Fig. 2 shows the architecture of the proposed *CNN*.

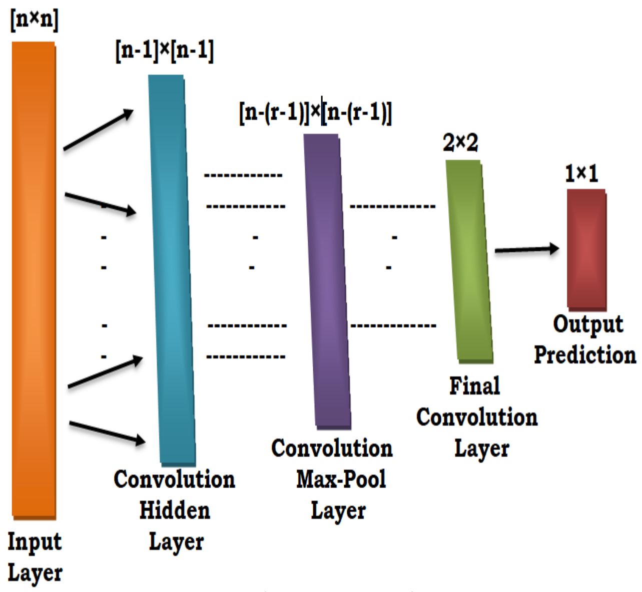

Figure 2: Proposed CNN Classifier Architecture for Heart Disease Data Sets

CNN is one kind of deep neural network. It can be study concurrently. The proposed detailed analysis of the process of *Convolutional Neural Networks algorithm* both the forward process and back propagation. Then applied the particular *Convolutional Neural Network* to implement the typical heart disease dataset problems. The bellow Fig. 3 shows general architecture of the *CNN*.

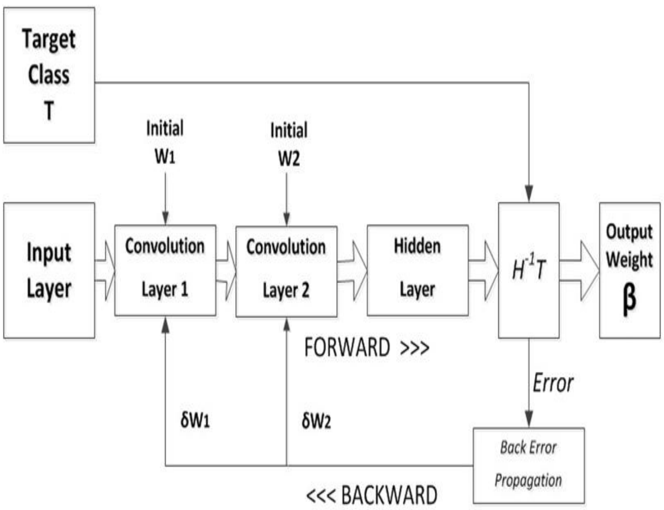

Figure 3: General Convolutional Neural Network Architecture

In addition, by measuring the actual time of forward and backward évaluation, analyse the maximal speed up and parallel efficiency theoretically.

4. 1 Role of Convolutional Neural Networks

In general, the structure of *CNN* includes two layers, one is feature extraction layer, and the input of every neuron is connected to the local except fields of the previous layer, and extracts the local feature. Once the local features are extracted, the positional relationship between it and other features will be evaluated. The other layer is feature map layer; all calculation layers of the network is mapped a multitude in feature map. Every feature is a level, and the weights of the neurons in the level are close.

The construction of feature applies the sigmoid function as stimulation function of the convolution network, which makes the feature shift in-variance. The numeral of free parameters of the network is reduced. All convolution layers in *CNN* by calculating the layer which is applied to calculate the local mean and the second extract. This particular two feature extraction constructions reduce the record's matrix level size. Multi-dimensional input vector of heart disease data sets can exit the network, which avoids the quality of data reconstruction in feature extraction and classification procedure.

4. 2 Feature Selection Algorithm

Several feature ranking and feature selection algorithms have been projected in the machine learning study. The purpose of these algorithms identifying and locate the unfit or unnecessary of from the vector. For the implementation of this research work, it applied feature ranking and selection modes by two initial steps of overall architecture: subset creation and subset calculation for the ranking of all features in every data set. Filter mode was applied to measure all subsets.

4. 3 Information Gain

The proposed feature choice both class membership and the presence/absence of a specific period are observed as random variables and one evaluates how more information around the class membership is increased by finding the presence/absence statistics as applied in decision tree induction. So, if the class membership is taken as a random variable C with two values, positive and negative, and a word is similarly observed as a random variable 'T' with two values, present and absent, it is by applying the information practical statement of ordinary information specified as the following equation:

$$IG(T) = H(C)\ H(C/T) = \Sigma\tau,\ cP(C=c,\ T=\tau)\ In\ [cP(C=c,T=\tau)/cP(C=c).P(T=\tau)] \quad (1)$$

Here, τ ranges over {present, absent} and c ranges over {c+, c^{-}}. As pointed out above, this is the measure of information about C (the class label) that increases by finding 'T' (presence or absence of a word).

4. 4 Back Propagation Algorithm

Convolutional Neural Networks technique is a multilayer perceptron, which is the specific system for recognition of two-dimensional data. It has many layers: input layer, convolution layer, sample layer and output layer. The *CNN* has two primary procedures: convolution and sampling. Convolution procedure applies a predictable filter F_x, re-convolution of the input data (the initial phase is the input data, the input of the later convolution is the feature data of every layer, namely, Feature Compose). Then, add a bias b_x, get convolution b_x layer C_x. A selecting procedure: n points of each neighborhood by pooling steps, get a point, and then by scalar weighting W_x+1 weighted, add bias b_x+1, and then by an activity function, make a narrow n times feature S_x+1.

The central engineering of *CNN* is the local tract, jointing of weights, sub selecting by time or space, and hence the training parameters extract feature and reduce the size. The benefit of *CNN technique* is prevention of explicit feature extraction, and learning from the training data. The neuron weights on the surface of the feature composing, thus, the network can see parallels, and reduce the multilevel of the network, Adapting sub sampling structure by the time or space, can attain a few degrees of robustness, scale and modification replacement. Input information and network topology can be a very good match, It has specific benefits in speech identification and data sets processing.

$$O_{x,y}^{1,k} = \tanh \sum_{t=0}^{f-1} \sum_{r=0}^{k_h} \sum_{c=0}^{k_w} W_{(r,c)}^{(k,t)}\, O_{(x+r,\ x+c)}^{(l-1,\ x)} + Bias^{(l,k)} \quad (2)$$

Among them, f is the numeral of convolution cores in a feature pattern, output of neuron of row x, column y in the l^{th} sub sample layer and k^{th} feature pattern:

$$O_{x,y}^{1,k} = \tanh\left(W^k \sum_{r=0}^{s_h} \sum_{c=0}^{s_w} O_{(x\times s_h+r,\ y\times s_w+c)}^{l-1,\ t} + Bias^{(l,k)} \right) \quad (3)$$

The output of the j^{th} neuron in l^{th} hides layer H:

$$O_{(i,j)} = \tanh\left(W^k \sum_{k=0}^{S-1} \sum_{x=0}^{S} \sum_{y=0}^{S_w} W_{(x,y)}^{(j,k)}\, O_{(x,y)}^{(l-1,k)} + Bias^{(l,k)} \right) \quad (4)$$

Among them, s is the number of feature patterns in sample layer output of the i^{th} neuron l^{th} output layer F:

$$O_{(i,j)} = \tanh\left(\sum_{j=0}^{H} O_{(l-1,j)} W_{(i,j)}^{l} + Bias^{(l,i)} \right) \quad (5)$$

4. 5 Modified Back-Propagation

A speed matrix depends on the technique to evaluate the output from a Neural Network (NN). Especially, it is an excellent mode of acquiring the notation used in back-propagation. Back-propagation is a Neural Network learning algorithm. The neural networks field was addressed by psychologists and neuro-biologists who wanted to create and test evaluation analogy of neurons. An NN is a set of input/output units in which every attachment has a weight joined by it. During the learning stage, the network learns with changing the weights so far capable to predict the exact class label of the input tuples.

Output deviation of the k^{th} neuron in output layer O:

$$d(O_k^0) = y^k - \tau^k \quad (6)$$

Input deviation of the k^{th} neuron in output layer:

$$d(I_k^0) = (y^k - \tau^k)\ \varphi(v_k)\ d(O_k^0) \quad (7)$$

Weight and bias variation of k^{th} neuron in output O:

$$\Delta W_{k,x}^0 = d(I_k^0)\ y_{k,x} \quad (8)$$

$$\Delta Bias_k^0 = d(I_k^0) \quad (9)$$

Output bias of k^{th} neuron in hide layer H:

$$d(O_k^H) = \sum_{i=0}^{i++} d(I_k^0) W_{i,k} \quad (10)$$

Input bias of k^{th} neuron in hide layer H:

$$d(I_k^H) = \varphi(v_k)\ d(O_k^H) \quad (11)$$

Where, weight and bias variation in row x , column y, in the m^{th} feature pattern, a previous layer in front of k neurons in hide layer H.

$$\Delta W_{m,x,y}^{H,k} = d(I_k^H)\ y_{x,y}^m \quad (12)$$

$$\Delta Bias_k^H = d(I_k^H) \quad (13)$$

Output bias of row x, column y in m^{th} feature pattern, sub sample layer S:

$$d(O_{x,y}^{S,m}) = \sum_k^{i++} d(I_{m,x,y}^H) W_{m,x,y}^{H,k} \quad (14)$$

Input bias of row x, column y, in m^{th} feature pattern, sub sample layer S:

$$d(I_{x,y}^{S,m}) = \varphi(v_k)\ d(O_{x,y}^{S,m}) \quad (15)$$

Weight and bias variation of row x, column y, in m^{th} feature of sub sample layer S:

$$\Delta W^{S,m} = \sum_{x=0}^{fh} \sum_{y=0}^{w} d(I_{\frac{x}{2},\frac{y}{2}}^{S,m})\ O_{x,y}^{C,m} \quad (16)$$

Among them, C represents convolution layer:

$$\Delta Bias^{S,m} = \sum_{x=0}^{fh} \sum_{y=0}^{fw} d\ (O_{x,y}^{S,m}) \quad (17)$$

Output bias of row x, column y in k^{th} feature pattern, convolution layer C:

$$d(O_{x,y}^{C,k}) = d(I_{\frac{x}{2},\frac{y}{3}}^{S,k})\ W^k \quad (18)$$

Input bias of row x, column y in k^{th} feature pattern, convolution layer C:

$$d(I_{x,y}^{C,k}) = \varphi(v_k)\ d(O_{x,y}^{C,k}) \quad (19)$$

Weight variation of row r, column c in m^{th} convolution core, corresponding to k^{th} feature pattern in l^{th} layer, convolution C:

$$\Delta W_{r,c}^{k,m} = \sum_{x=0}^{fh} \sum_{y=0}^{fw} d(I_{x,y}^{C,k})\ O_{x+r,y+c}^{l-1,m} \quad (20)$$

Total bias variation of the convolution core:

$$\Delta Bias^{C,k} = \sum_{x=0}^{fh} \sum_{y=0}^{fw} d\ (I_{x,y}^{C,k}) \quad (21)$$

4. 6 CNN Algorithm Process Steps

Step 1: Take the arff data set.
Step 2: Feature selection using information gain and ranking.
Step 3: Classification algorithm.
Step 4: Every Feature calculates f_x value of input layer.
Step 5: Bias class of every feature calculation.
Step 6: After giving the feature map, it goes to forward pass input layer.
Step 7: Evaluate the convolution cores in a feature pattern.
Step 8: Gives sub sample layer and feature value.
Step 9: Back-propagation input deviation of the k^{th} neuron in output layer.
Step 10: Lastly, produce the chosen feature and classification results of (1*1) matrix output (i.e. Normal or Abnormal of heart disease data set records).

5. Results and Discussion

The dataset of all records having a set of 13 attributes. During the collection of a few missing values are found because of manual write down by doctors. By means of data missing Weka 3.6.6 tool the missing values are replaced using mean, mode method by filter option available in the tool. A confusion matrix is obtained to compute the accuracy of classification. In this work, confusion matrix shows how many instances have been assigned to every class. The bellow Table 2 shows the confusion matrix for evaluation of performance measures.

Table 2: Class A = YES (has heart disease); Class B = NO (No heart disease)

	A (has heart disease)	**B (no heart disease)**
A (has heart disease)	True Positive (TP)	False Negative (FN)
B (no heart disease)	False Positive (FP)	True Negative (TN)

The data set made of 573 records and it has been divided into 25 classes, where each class consists of 23 records. Training and Testing sets are divided into equal amount to predict the exactness of the system. To measured certainty of 13 attributes are chosen to classify the system by a symbolic learning approach by interval method, i.e., (μ-σ, μ+σ) as indicate in the Table 1.

In this work trained the classifiers to classify the heart diseases are either ''has heart disease'' or ''no heart disease''. The general and particular of confusion matrixes are two classes (i.e. normal and abnormal) of five classifiers. The bellow Table 3 shows for the evaluation of True Positive, True Negative, False Positive, and False Negative values in confusion matrix.

Table 3: Confusion matrix accès frome *CNN* classifier

	A	**B**
A	257	05
B	08	271

The proposed *CNN algorithm* compared with the classifiers such as *K-NN*, *Naïve Bayes*, *SVM, and ANN*, and evaluated their individual Confusion Matrix respectively.

5. 1 Accuracy, Mean Value Error

The classifiers of *Accuracy* and *Mean Value Error* performance measures are evaluated by using below formulae and results are shown in the bellow Table 4, and the comparison chart is specified in the bellow Figure 4.

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

$$\text{Mean Value Error} = \frac{1}{n}\sum \lfloor y_i - x_i \rfloor$$

Where: TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative

2nd equation MVE: y_i is actual value of ith observation, x_i *is* calculated value of ith observation; n is total number of observations.

Table 4: Accuracy, and Mean Value Error comparative of five classifiers

Classifiers	**Accuracy**	**Mean Value Error**
K-NN	54.42	0.046
Naive Bayes	75.68	0.003
SVM	79.58	0.009
ANN	87.87	0.0028
CNN	97.43	0.0014

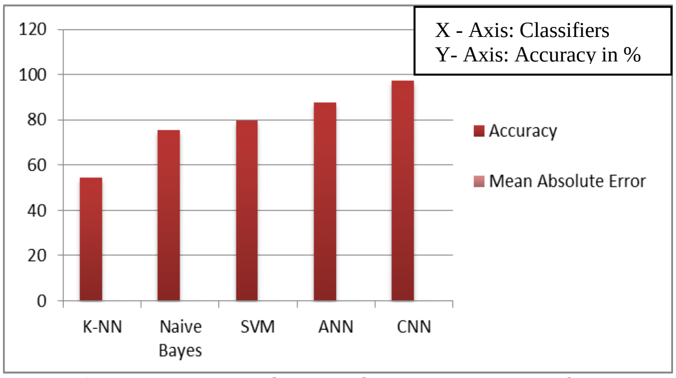

Figure 4: Accuracy and Mean Value Error Comparison Chart

Similarly, the *Overall accuracy* performance measure evaluated for data set among unreliable K values shows in the bellow Table 5, and the comparison chart is specified in the bellow Figure 5.

Table 5: Overall Accuracy accès for data set among unreliable K values

K	**K-NN**	**Naïve Bayes**	**SVM**	**ANN**	**CNN**
1	0.56	0.77	0.79	0.96	0.98
3	0.52	0.77	0.75	.93	0.99
5	0.50	0.74	0.73	0.88	0.95
7	0.48	0.71	0.71	0.84	0.94
9	0.45	0.70	0.79	0.81	0.88

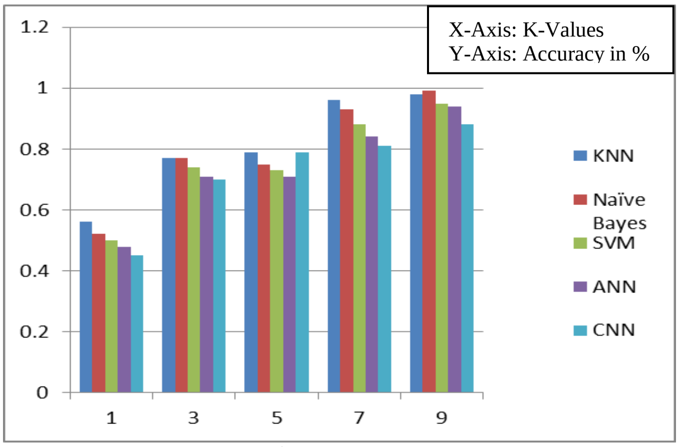

Figure 5: Comparaison chart for Overall Accuracy access data set among unreliable K values

5. 2 Sensitivity, Specificity, and F-Measure

The classifiers of *Sensitivity, Specificity, and F-Measure* performance measures are evaluated by using bellow formulae and results are shown in the bellow Table 6, and their comparison chart is specified in the bellow Figure 6.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

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

$$\text{F} - \text{Measure} = \frac{2 \times \text{Sensitivity} \times \text{Specificity}}{\text{Sensitivity} + \text{Specificity}}$$

Table 6: Comparison Table for Sensitivity, Specificity, and F-Measure

Classifiers	Sensitivity	Specificity	F-Measure
K-NN	0.9467	0.9143	0.8520
Naive Bayes	0.9507	0.9421	0.9422
SVM	0.9534	0.9435	0.9455
ANN	0.9638	0.9569	0.9521
CNN	0.9645	0.9577	0.9532

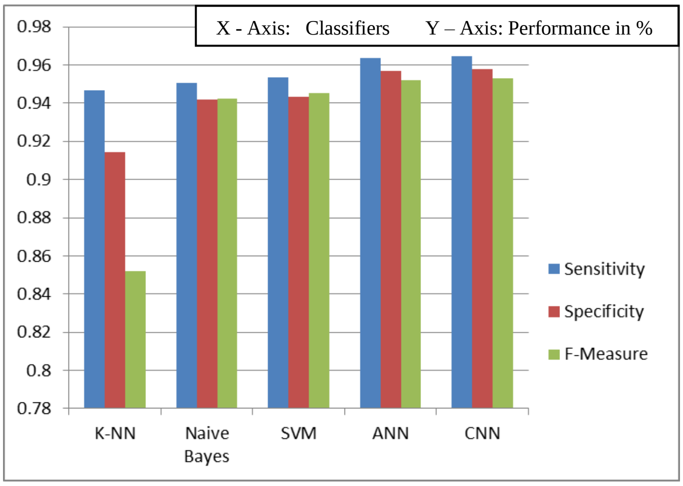

Figure 6: Comparison chart for Sensitivity, Specificity, and F-Measure

5. 3 False Positive Rate, False Negative Rate and Overall Error Rate

The classifiers of *False Positive Rate, False Negative Rate and Overall Error Rate* performance measures are evaluated by using bellow formulas and results are shown in the Table 7, and their comparison chart is specified in the Figure 7.

$$\text{False Positive Rate} = \frac{FP}{FP+TN}$$

$$\text{False Negative Rate} = \frac{FN}{FN + TP}$$

$$\text{Mean Value Error} = \frac{1}{n}\sum \lfloor y_i - x_i \rfloor$$

Table 7: Comparison Table for FPR, FNR, and Overall Error Rate

Classifiers	FPR	FNR	Overall Error Rate
K-NN	0.0631	0.0596	0.0678
Naive Bayes	0.0568	0.0382	0.0453
SVM	0.0487	0.0356	0.0417
ANN	0.0321	0.0267	0.0345
CNN	0.0398	0.0298	0.0343

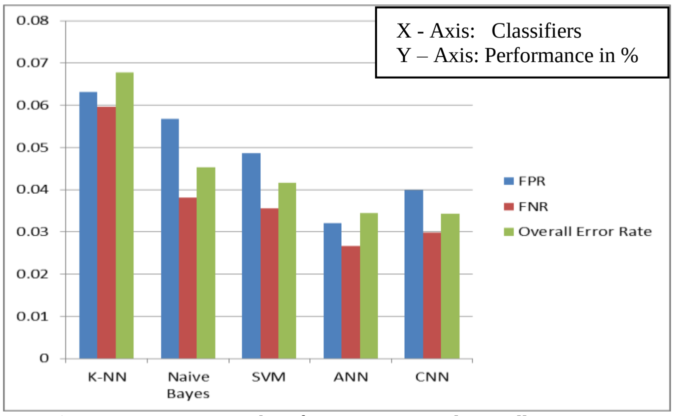

Figure 7 : Comparison chart for FPR, FNR, and Overall Error Rate

The evaluated measurement of the performance based on 13 attributes are accurate in natural history, and the accuracy of *Convolution Neural Network* using heart disease data set is reached up to 99%. C*NN* is improved performance than other classifiers, while *K-NN* shows weak performance than other classifiers.

6. Conclusion and Future Scope

In this proposed system, *Convolution Neural Network* classification methodology is made efficiently. To perform classification task of medical data, the neural network is trained using Convolution technique. The experimentation is shown with heart disease dataset by considering the single and multilayer neural network methods. *Convolution Neural Network* technique is a multilayer perceptron that is the superior design for reorganization of two-dimensional heart disease data information. In this so many layers are observed, i.e.: input layer, convolution layer, sample layer and output layer. In addition, in deep neural network architecture the convolution layer and sample layer can have multiple layers in between. The overall objective of this work is to predict the heart diseases with more accuracy. This work can be enhanced by increasing the more number of records with attributes to provide better accuracy for the existing system.

Data Availability

1. Clevelanddatabase:http://archive.ics.uci.edu/ml/datasets/Heart disease
2. Stalogdatabase: https://archive.ics.uci.edu/ml/datasets/statlog+(heart)

Acknowledgements

All sincerely, would like to express my gratitude to both, the management and department of Computer Science and Engineering of *Neil Gogte Institute of Technology (NGIT).*

References

[1] Waseem Rawat, Zenghui Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review", Neural Computation © 2017 Massachusetts Institute of Technology, Vol.**29**, Issue.**9**, pp.**24-36**, September **2017.**

[2] Hafsa Binte Kibria, et. al., "The Severity Prediction of the Binary and Multi - Class Cardiovascular Disease - A Machine Learning-Based Fusion Approach", ar Xiv: 2203.04921v1 [cs. LG], A PREPRINT - March 10th, pp.**1-20, 2022.**

[3] Tsehay Admassu Assegie, "Heart disease prediction model with k-nearest neighbor algorithm", International Journal of Informatics and Communication Technology (IJ-ICT), ISSN: 2252-8776, DOI: 10.11591/ijict.v10i3, Vol.**10**, No.**3**, pp.**225-230**, December **2021.**

[4] An Dinh, Stacey Miertschin, Amber Young, and Somya D Mohanty. A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. BMC medical informatics and decision making, Vol.**19**, Issue.**1**, pp.**2-11**, **2019.**

[5] Archana Singh and Rakesh Kumar. Heart disease prediction using machine learning algorithms. In 2020 international conference on electrical and electronics engineering (ICE3), IEEE, pp.**452–457, 2020.**

[6] Md. Mamun Ali, Bikash Kumar Paul, Kawsar Ahmed, FrancisMBui, JulianMWQuinn, and Mohammad Ali Moni. Heart disease prediction using supervised machine learning algorithms:

performance analysis and comparison. Computers in Biology and Medicine, 136:104672, **2021.**

[7] Devansh Shah, Samir Patel, and Santosh Kumar Bharti. Heart disease prediction using machine learning techniques. SN Computer Science, 1(6):1–6, pp.**244-252, 2020.**

[8] Manoj Diwakar, Amrendra Tripathi, Kapil Joshi, Minakshi Memoria, Prabhishek Singh, et al. Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37: pp.**3213–3218, 2021.**

[9] Hafsa Binte Kibria, Abdul Matin, Nusrat Jahan, and Sanzida Islam. A comparative study with different machine learning algorithms for diabetes disease prediction. In 2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), IEEE, pp.**1–8**, **2021.**

[10] Aleh Ayatollahi, Leila Gholamhosseini, and Masoud Salehi. Predicting coronary artery disease: a comparison between two data mining algorithms. BMC public health, Vol.**19**, Issue.**1**, pp.**1–9, 2019.**

[11] Zeinab Arabasadi, Roohallah Alizadehsani, Mohamad Roshanzamir, Hossein Moosaei, and Ali Asghar Yarifard. Computer aided decision making for heart disease detection using hybrid neural network-genetic algorithm. Computer methods and programs in biomedicine, 141:pp.**19–26, 2017.**

[12] Aditi Gavhane, Gouthami Kokkula, Isha Pandya, and Kailas Devadkar. Prediction of heart disease using machine learning. In 2018 second international conference on electronics, communication and aerospace technology (ICECA). IEEE, 2018. 18 A PREPRINT - MARCH 10, pp.**1275–1278**, **2022.**

[13] Monther Tarawneh and Ossama Embarak. Hybrid approach for heart disease prediction using data mining techniques. In International Conference on Emerging Internetworking, Data & Web Technologies, Springer, pp.**447–454, 2019.**

[14] Saman Iftikhar, Kiran Fatima, Amjad Rehman, Abdulaziz S Almazyad, and Tanzila Saba. An evolution based hybrid approach for heart diseases classification and associated risk factors identification. Biomedical Research, Vol.**28**, Issue.**8**, pp.**3451–3455, 2017.**

[15] Ninad Marathe, et., al., "Prediction of Heart Disease and Diabetes Using Naive Bayes Algorithm", International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, doi.org/10.32628/CSEIT217399, Vol.**7**, Issue.**3**, pp.**447-453**, 31 May **2021.**

[16] Yamala Sandhya, "Prediction of Heart Diseases using Support Vector Machine", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; Vol.8, Issue.**11**, pp.**126-135**, Feb' **2020.**

AUTHORS PROFILE

Dr. V. Krishnaiah having 22 years of teaching experience in reputed Engineering Institutions and presently working as Associate Professor, Dept. of CSE, *Neil Gogte Institute of Technology*, Affiliated by Osmania University, Hyderabad, Telangana State, India. I received Ph.D. (Computer Science and Engineering), and M.Tech. (Computer Science) both from College of Engineering (CEH), Jawaharlal Nehru Technological University Hyderabad (JNTUH), Hyderabad, Telangana State, India; M.Phil. (Mathematics) from Alagappa University (AU), Karaikudi, Tamilnadu State, India; and M.Sc. (Mathematics) from Osmania University (OU), Hyderabad, Telangana State, India.

Published various publications in International/National Journals and also presented papers in National /International conferences.

Google Scholar: Citations – **600+**, H Index – **06;** i10 Index – **06**.
https://scholar.google.co.in/citations?hl=en&user=mP3O0AYAAAAJ

My research interests include Data Mining, Network Security, Image Processing, Machine Learning, and Data Science, AIML.

Awarded as "Certificate of Excellence in Reviewing", 'In recognition of an outstanding contribution to the quality of journal", from 'South Asian Journal of Social Studies and Economics' (SAJSSE), UK, Landon.

