IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.2, February 2021 E-ISSN: 2347-2693

Performance Analysis of Convolutional Network System for Heart Disease Prediction

Julie M. David[1*], Sarika S.[2]

[1,2]Department of Computer Applications MES College Marampally, India

**Corresponding Author: julieeldhosem@yahoo.com, Tel.: +91 9447104152*

DOI: https://doi.org/10.26438/ijcse/v9i2.1722 | Available online at: www.ijcseonline.org

Received: 15/Feb/2021, Accepted: 18/Feb/2021, Published: 28/Feb/2021

Abstract **-** Heart is one of the major parts of human body, which maintains life line. It pumps the blood and supplies to all parts of the body. Heart disease prediction is significant work. Here we propose a Heart disease prediction model and is a hybrid intelligent system developed using classifier such as deep learning, feature extraction tools, and normalization methods. This intelligent system shows the high accuracy than the other datamining classifier. In this paper, the proposed model will help the medical field to reduces the cost and work load, and also ensures the accuracy of result. This System is very efficient and effective**.**

Keywords **-** Heart disease, Deep learning, Classification

I. INTRODUCTION

Data mining is used for finding patterns and regularities from a huge data by using the techniques such as data analysis and software techniques. Nowadays, data mining plays vital role in predicting multiple disease. By using data mining techniques, the number of tests can be reduced [1]. Large number of people die of heart disease annually [2].Normally, machine learning model has to be trained using a massive dataset, which contains a huge number of parameters, which requires a high computation task. To analyse large amount of data, machine learning techniques are most useful [3].Heart pumps blood through the blood vessels [4]. In traditional, a number of medical tests are required to predict the disease. Quality of service is the major problem in healthcare industry [5]. If the data mining and machine learning techniques will be integrated with the medical information system, then it will be highly advantageous [6]. The advantage is to help a non-specialized person to make an intelligent decision about the risk level in heart disease [7]. Several different symptoms are associated with heart disease, which makes it difficult to diagnose it quicker and better [8]. Neural network is a commonly used technique for predicting heart disease. In this paper, a Performance Analysis of Convolutional network System for the Heart disease Prediction is developed. There are number of factors which increases the risk of heart, they are; Blood pressure, cholesterol, family background of heart disease, lack of exercise etc. Common cardiovascular diseases include coronary heart disease, cardiomyopathy, hypertensive heard disease, heart failure, etc [9]. Heart disease prediction is performed in order to reduce the heart disease risk and signs, symptoms, physical examination of a patient etc. are used for disease prediction. Almost all the medical practitioners predicting the disease by learning and by their experience but these leads to a number of difficulties such as less accurate results, less experience, time dependent performance etc listed in following figure [10]. To improve the prediction accuracy and to reduce the prediction time, neural networks are used. Neural networks are the best tool for prediction of diseases [11]. This work is focused on the Performance Analysis of Convolutional network System for the Heart disease Prediction by using python and Rapid Miner tool.

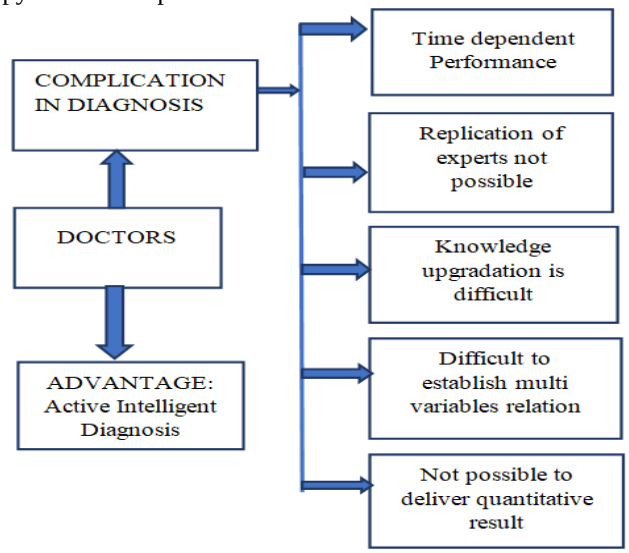

Figure.1: Complexity in diagnosis with doctor

Section I contains the introduction of Performance Analysis of Convolutional Network System For The Heart Disease Prediction , Section II contain the related work of Performance Analysis of Convolutional Network System For The Heart Disease Prediction , Section III contain the some measures of Performance Analysis of Convolutional Network System For The Heart Disease Prediction, Section IV contain the architecture and essential steps of

Performance Analysis of Convolutional Network System For The Heart Disease Prediction, section V explain the Performance Analysis of Convolutional Network System For The Heart Disease Prediction methodology with flow chart, Section VI describes results and discussion of Performance Analysis of Convolutional Network System For The Heart Disease Prediction, Section VII contain the recommendation of Performance Analysis of Convolutional Network System For The Heart Disease Prediction and Section VIII concludes research work with future directions).

II. RELATED WORKS

A survey is carried out on different data mining and machine learning techniques and the different parameters used for heart disease prediction are discussed.
K.Gomathi, D. Shanmaya Priya [1] suggested a method called multi disease prediction. And the prediction is done by using data mining techniques. In this paper, two different data mining classification techniques (Naïve Bayes and J48) are used and their performance was compared in order to evaluate the best classifier.

Boshra Bahrami et al. [2] developed various data mining techniques to evaluate the prediction and diagnosis of heart disease. The main objective of this paper is to evaluate different classification techniques such as J48, decision tree, KNN, naïve bayes, and SMO. The comparison results show that J48 decision tree is the best classifier.

Ramalingam et al. [3] was suggested a model for predicting Heart disease using machine learning techniques. The model is based on supervised learning algorithm. This paper provides a survey of various models based on such algorithms and techniques and analyzes their performance.

Entin Martiana Kusumaningtyas, et al. developed a model for Feature Extraction for Application of Heart Abnormalities Detection [4]. This model is proposed a new method of iridology and computation process to find the heart problems. The extraction of features results in high performance.

Sairabi H Mujwar, P.R. Devale [5] used k-means and Naïve Bayes to predict heart disease. In this paper an improved k-means algorithm is proposed which tries to remove one of the major limitations of basic k-means algorithm, which requires number of clusters as input.

Abhishek Taneja [6] developed a model using data mining technique to predict heart disease and the model is focused on data mining techniques from trans thoracic ecg report dataset. This allows to improve reliability of heart diagnosis using ECG. The efficiency of the model was evaluated using the standard metrics of accuracy.

An intelligent system for heart disease prediction is developed by Purushottam et al. [7]. The system is developed using data mining techniques. 10-fold method is used to train and test the system. After applying 10-fold method, the accuracy in testing and training phase are 86.3% and 87.3% respectively.

Jaymin Patel and Dr. Samir Patel [8] suggested data mining and machine learning techniques to predict heart disease. In this paper, it shows that J48 tree technique is the best classifier and it provides more accuracy and least total time to build. J48 algorithm based on UCI data has the accuracy of 56.76% and building time is 0.04 seconds.

By focusing on SVM - Decision trees - Logistic regression Mythili T et al. [9] developed a heart disease prediction model. The objective of the developed model is to proposes a predictive framework for disease and also compare the efficiency of the model by merging the outcomes of multiple models using a single model.

Effective Heart Disease Prediction using Hybrid Machine Learning Techniques [11] has introduced by Senthil Kumar Mohan, Chandra Segar Thirumalai and Gautham Srivaastava. This paper aims to find out significant features by using technique such as machine learning. It produces an improved accuracy of 88.7% through Hybrid Random Forest with Linear Model.

Syed Umar Amin, Kavita Agarwal, Rizwan Beg [12] proposed a method called Genetic neural network based data mining in prediction of heart disease using risk factors. In this paper, major risk factors are used to predict heart disease. This method uses two data mining tools, neural network and genetic algorithms. The advantage of the hybrid method is fast, more stable and accurate as compared to back propagation. The system was implemented in Matlab and predicts the heart disease risk with an accuracy of 89%.

Using classification-based data mining techniques, Sujata joshiet al. [13] proposed a method for predicting heart disease. The results show that KNN has highest accuracy. But The decision tree performs well when compared to other two methods when used for prediction.

Sellappan Palaniappan and Rafiah Awang [14] focuses on data mining techniques that can predict heart disease. The system is developed using three data mining classification modelling techniques. Naïve Bayes algorithm followed by the algorithms neural network and decision trees is the most efficient model to predict heart disease.

Model for predicting Heart Disease has developed by Hlaudiet al. [15] using Classification Algorithms. Data mining algorithms such as CART, J48, Naïve Bayes, REPTREE, and Bayes Net are applied in this research for predicting heart attacks. The prediction accuracy is 99%.

G. Parthiban et al. [16] has developed a model for heart disease for diabetic patients. Model is developed by applying methods in machine learning. In this paper they have tried to predict the chances of getting a heart disease using attributes from diabetic's diagnosis and shown that it is possible to diagnose heart disease vulnerability in diabetic patients with reasonable accuracy.

III. METHODOLOGY

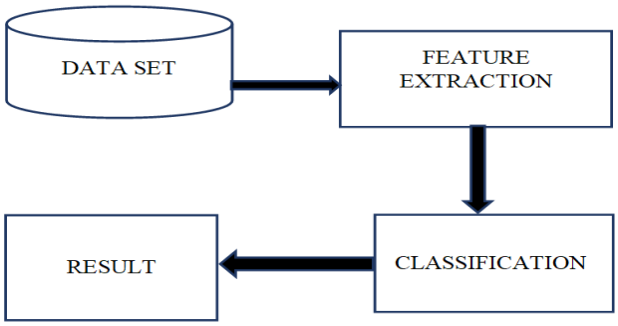

Figure.2: Flow chart

A. DATA SET

Data set is a collection of data. In this, we use heart disease data set and it is available publicly and this database contains 14 attributes, the target field in the database is to determine the presence of heart disease in the patient. It is denoted by integer values 0 and 1. In which 0 refers to the absence and 1 represents the presence of disease. The data set using in this model is shown in table 1. The experiment is carried out on a dataset which is publicly available. The dataset contains total304 records. A data mining tool Rapid miner is used for experiment. Sample of dataset is listed below.

Table 1: Data set

Sl No.	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
1	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
2	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
3	40	1	0	110	167	0	0	114	1	2	1	0	3	0
4	60	1	0	117	230	1	1	160	1	1.4	2	2	3	0

Table 2: Description of parameters used in dataset

Sr. no	Attribute
1	Age(In Years)
2	Sex (Male -1 And Female-0)
3	Cp (Type of Chest Pain)
4	Thestbps (Resting Bp)
5	Chol (Serum Cholesterol)
6	Restecg (Resting Electrographic Result)
7	Fbs (Fasting Blood Sugar)
8	Thalach(Max. Heart Rate Achieved)
9	Exang(Exercise Induced Angina)
10	Oldpeak (St Depression Induced by Exercise Releative To Rest)
11	Slope (Slope 0f The Peak Exercise ST Segment)
12	Ca (Number of Major Vessels Colored by Floursopy)
13	Thal (Defect type)
14	Target (Presence or absence of disease (1- presence of disease and 0-absence of disease))

B. FEATURE EXTRACTION

Nowadays it is becoming quite common to be working with large datasets that consists of large number of features. Feature extraction is process of reducing the dimensionality to manageable size. These extracted features are used for processing the data. The extracted features are used to accurately and completely describing the original data set.

Using rapid miner tool, we extract the features using the extract aggregates operator. One of the important advantages of this is that the resulting feature space is homogeneous. Extract aggregate operator calculate descriptive features and these features are later provided at the output port for further operation. This operator calculates some features of data for each item. One of the important advantages is that the resulting feature space is homogeneous and the calculated features are used as the input of the inner sub process. The result of feature extraction is shown in below table.

Table 3: Feature extraction result using extract aggregates

Row No.	Time series	Sum	Mean	Geometric mean	First quartile	Median	Third quartile	Min	Max	Std deviation	Kurtosis	skewness
1	age	16473.0	54.366	53.569	47.0	55.0	61.0	29.0	77.0	9.082	-0.542	-0.202
2	sex	207.0	0.683	0.0	0.0	1.0	1.0	0.0	1.0	0.466	-1.382	-0.791
3	cp	293.0	0.966	0.0	0.0	1.0	2.0	0.0	3.0	1.032	-1.193	0.484
4	trestbps	39882.0	131.623	130.502	120.0	130.0	140.0	94.0	200.0	17.538	0.929	0.713
5	chol	74618.0	246.264	241.172	211.0	240.0	275.0	126.0	564.0	51.830	4.505	1.143
6	fbs	45.0	0.148	0.0	0.0	0.0	0.0	0.0	1.0	0.356	1.959	1.986
7	restecg	160.0	0.528	0.0	0.0	1.0	1.0	0.0	2.0	0.525	-1.362	0.162

8	thalach	45343.0	149.646	147.740	133.0	153.0	166.0	71.0	202.0	22.905	-0.061	-0.537
9	exang	99.0	0.326	0.0	0.0	0.0	1.0	0.0	1.0	0.469	-1.458	0.742
10	oldpeak	315.0	1.039	0.0	0.0	0.8	1.6	0.0	6.2	1.161	1.575	1.269
11	slope	424.0	1.399	0.0	1.0	1.0	2.0	0.0	2.0	0.616	-0.627	-0.508
12	ca	221.0	0.729	0.0	0.0	0.0	1.0	0.0	4.0	1.022	0.839	1.310
13	thal	701.0	2.313	0.0	2.0	2.0	3.0	0.0	3.0	0.612	0.297	-0.476
14	target	165.0	0.544	0.0	0.0	1.0	1.0	0.0	1.0	0.498	-1.980	-0.179

Using the rapid miner auto model, we can extract 6 attributes from the original dataset. They are as shown in Table 4.

Table 4: Extracted features

Sl No.	sex	cp	trestbps	slope	ca	thal
1	1	3	145	0	0	1
2	1	2	130	0	0	2
3	1	0	110	1	0	3
4	1	0	117	2	2	3

C. CLASSIFICATION

The classification concept concerned with building a model and constructed model groups data in the dataset into various classes.Classification is to determine the class label for a given data. The objective of classification is to determine the class label of unknown data effectively. Some of the benefits of using deep learning approach are - no need for feature engineering, elimination of unnecessary costs, efficient at delivering high quality results, etc.

Here, deep learning classification algorithm is used. The techniques and functions of deep learning is similar to machine learning. However, the capabilities are different. A layered structure of algorithm is used to achieve the prediction accuracy and these layers are able to learn and make effective decisions on its own.

D. NEURAL NETWORK BASED HEART DISEASE PREDICTION

In this paper, the system is developed for Performance Analysis of Convolutional network System for the Heart disease Prediction. Historical heart disease dataset is used and the system uses 14 attributes such as age, sex, BP, etc. These attributes are important for the disease prediction. The technique is to develop a convolutional neural network.

E. CONVOLUTIONAL NEURAL NETWORK

CNNs have commonly used in classification methods to achieving higher accuracy over large datasets. It is because of their ability to learn rich features from raw data as well as classifier learning. Main operations of a typical CNN architecture is shown in Figure 3. [17].

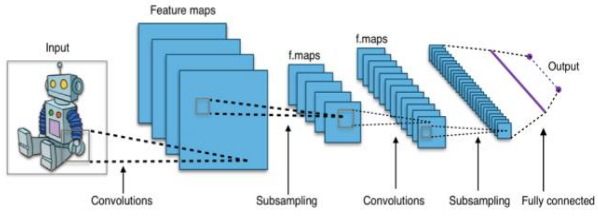

Figure 3: Main operations of a typical CNN architecture

In convolution operation, several different size kernels are applied on the input data to generate feature maps. These generated feature maps are then considered as the input to the next operation and where maximum activations are selected from them within small neighbourhood. These operations are useful for reducing the dimensions of feature vectors. In fully connected layer, high level abstractions are modelled and the weights of neurons are updated for better performance of input data.

H2O's deep learning algorithm is based on feed forward artificial neural network. It is a multi-layer perceptron and it is trained with back-propagation using stochastic gradient descent. The network contains number of hidden layers consisting of neurons with activation functions and also advanced features like adaptive learning rate, momentum training etc.

H2O-3 is supported only by deep learning network. The algorithm is deterministic, if the reproducible parameter is true and, in this case, only one thread is used. In this model, number and size of each hidden layer and epoch value is "50, 50" and 10 respectively.

Table 5: Deep learning performance

Criterion	Value
Accuracy	85.1%
Classification_Error	14.9%
Auc	88.8%
Precision	82.1%
Recall	93.8%
F_Measure	87.4%
Sensitivity	93.8%
Specificity	74.6%

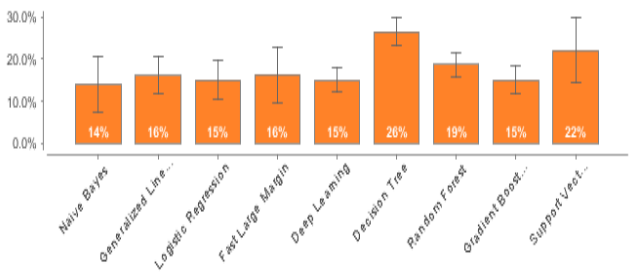

Figure 4: Classification error comparison

Table 6: Classification model comparison

MODEL	ACCURACY	STD. DEVIA-TION	GAINS	TOTAL TIME	TRAINING TIME	SCOR-ING TIME
Deep learning	85.1%	±2.9%	70	7s	2s	3 s
Decision tree	73.5%	±3.4 %	42	4s	224 ms	430 ms
Random forest	81.4%	±2.8 %	62	36s	828 ms	3 s
Gradient boosted tree	85.0%	±3.3%	74	55s	861 ms	190 ms
SVM	77.9%	±7.6%	56	14s	1s	339 ms

IV. RESULTS AND DISCUSSION

From the above results, it is clear that deep learning is an excellent method for classifying the data. Run time taken by deep learning model is less than Random forest, Gradient boosted Trees and Support Vector Machine. But the relative error of deep learning model is almost same as Generalized linear model, Random forest, Gradient boosted Trees and Support Vector Machine. Training time of deep learning model is 2s.

After applying neural networks on training data, the results obtained is represented in the form of a 2-D matrixes. This 2-D matrix is called, Confusion matrix. It compares the actual values with the predicted values [14].The effectiveness is determined using the matrix [15]. The confusion matrix is shown in table 7.

Table 7: Confusion matrix

	True no	**True yes**	**Class precision**
Predicted no	29	3	90.62%
Predicted yes	10	45	81.82%
Class recall	74.36%	93.75%	

A confusion matrix validates a particular classifier with a fixed threshold, But ROC curve validates the classifier over various possible thresholds.

From these results, 93.8% of the positives were successfully predicted by our model. So we can conclude that, the proposed model is an efficient model for the prediction.

Receiver Operating Characteristics (ROC) curve is to determine the performance of a multi-class classification problem over various thresholds. So, ROC is a probability curveie, curve is plotted with TPR against the FPR. Where y axis represents the TPR and x axis is FPR.

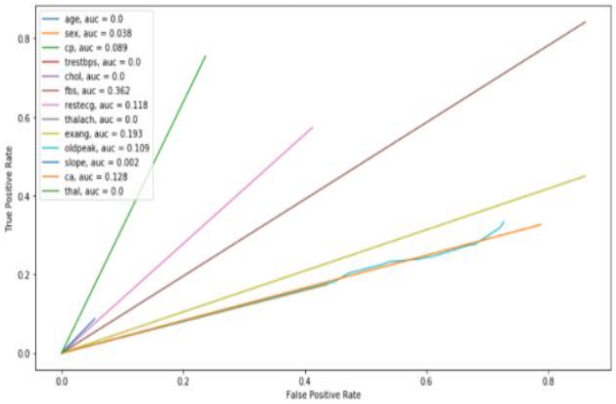

Figure 5: ROC Curve – Heart Disease Features

Method for classifying the data. Run time taken by deep learning model is less than Random forest, Gradient boosted Trees. But the classification error of deep learning model is almost same as Generalized linear model, Logistic Regression, Fast large margin, Random forest, Trees and deep learning. Training time of deep learning model is 2s. Original dataset contains 14 attributes. After feature extraction, the number of attributes reduced to 6. I.e., 6 attributes are enough to predict the model. The improvement of accuracy after feature extraction is from 81% to 85%. I.e., the accuracy before feature extraction is 81% and the accuracy after feature extraction is 85%. The accuracy of the proposed model is calculated as 85%. From these results, there is a small deviation in accuracy between before and after feature extraction. That is, accuracy of the model is greater after the feature extraction. The advantage of proposed model is less space and less time. Hence, we can also reduce the cost.

V. CONCLUSION

Nowadays heart disease is an important cause of death. The early prediction of heart disease is very important. From the study it shows that, techniques such as machine learning and also data mining are plays an essential role in heart disease classification. Here, we have studied various algorithms, techniques that can be used for classification of heart disease databases and also, we have seen the accuracy obtained by them. We propose a heart disease prediction model with improved accuracy.

This paper proposes a framework using hybrid algorithms to arrive at an accurate prediction of heart disease. This paper proposes a comparative study of the multiple results, which include sensitivity, specificity, and accuracy. So, the most effective and most weighed model can be found. The advantage of the proposed method is fast, more stable and accurate as compared to other methods. The system was implemented in Rapid miner tool and predicts the heart disease with an accuracy of 85%.

VI. FUTURE SCOPE

There are many possible scopes that could beexplored to improve the accuracy of thisprediction system. Due to time

limitation, the following work needs to be performed in the future. This paper proposes a framework using convolution network system to arrive at an accurate prediction of heart disease. Further work involves development of system using the methodology to be use for checking the imbalance with other data mining and machine learning models.

REFERENCES

[1] K.Gomathi, Dr. D.Shanmuga Priyaa, "*Multi Disease Prediction using Data Mining Techniques*", International Journal of System and Software Engineering, **pp.12-14, 2016.**

[2] Boshra Brahmi, Mirsaeid Hosseini Shirvani, "*Prediction and Diagnosis of Heart Disease by Data Mining Techniques*", Journals of Multidisciplinary Engineering Science and Technology, **Vol.2, pp.164168, 2015.**

[3] V.V. Ramalingam, Ayantan Dandapath, M Karthik Raja "*Heart disease prediction using machine learning techniques: A survey*", Int. J. of Engineering & Technology, **Vol.7, No. 2, pp684-687, 2018.**

[4] F. D. Kusuma, E. M. Kusumaningtyas, A. R. Barakbah and A. A. Hermawan, "*Heart Abnormalities Detection Through Iris Based on Mobile*," *In Proceedings of Int. Electronics Symposium on Knowledge Creation and Intelligent Computing*, Bali, Indonesia, **pp. 152-157, 2018.**

[5] Sairabi H.Mujawar, P.R.Devale, "*Prediction of Heart Disease using Modified K-means and by using Naïve Bayes*", Int. J. of Innovative research in Computer and Communication Engineering, **Vol.3, pp.10265-10273, 2015.**

[6] Abhishek Taneja, "*Heart disease Prediction System Using Data Mining Techniques*", Oriental Journal of Computer Science & Technology, ISSN: 0974-6471, **Vol. 6, No. 4, pp 457-466, 2013.**

[7] Purushottam, Prof. Dr. Kanak Saxena and Richa Sharma, "*Efficient Heart Disease Prediction System*", Procedia Computer Science 85, **pp.962-969, 2016.**

[8] Jaymin Patel, Prof.Tejal Upadhyay, Dr.Samir Patel, "*Heart Disease Prediction using Machine Learning and Data Mining Technique*", Int. J. of Computer Science and Communication, **pp.129-137, 2016**

[9] Mythili T, Dev Mukherji, Nikita Padalia, Abhiram Naidu, "*A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL)*", Int. J. of Computer Applications, 0975– 8887, **Vol.8, No.16, 2013.**

[10] K.S.Kavitha, K.V. Ramakrishnan, Manoj Kumar Singh, "*Modeling and design of evolutionary neural network for heart disease detection*", Int. J. of Computer Science Issues, **Vol. 7, No. 5, pp 272-283, 2010.**

[11] Senthilkumar Mohan, Chandrasegar Thirumalai, Gautham Srivastava, "*Effective Heart Disease Prediction using Hybrid Machine Learning Techniques*", IEEE Access,**Vol.7, pp 81542-81554, 2019.**

[12] S. U. Amin, K. Agarwal and R. Beg, "*Genetic neural network based data mining in prediction of heart disease using risk factors",* In Proceedings of *IEEE Conference on Information & Communication Technologies,* Thuckalay, India, **pp. 1227-1231, 2013.**

[13] Sujata Joshi, Mydhili K Nair, "*Prediction of heart disease using classification-based data mining techniques*", Computational Intelligence in Data Mining-. Springer, New Delhi, **Vol. 2, pp.503-511, 2015.**

[14] Sellappan Palaniappan and Rafiah Awang, "*Intelligent Heart Disease Prediction System Using Data Mining Techniques*", Int. J. of Computer Science and Network Security, **Vol. 8 No.8, 2008.**

[15] Hlaudi Daniel Masethe, Mosima Anna Masethe, "*Prediction of Heart Disease using Classification Algorithms*", In Proceedings of World Congress on Engineering and Computer Science, **Vol II, San Francisco, USA, 2014.**

[16] G. Parthiban, S.K. Srivatsa "*Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients",* Int. J. of Applied Information Systems, **Vol. 3, No.7, pp. 25-30, 2012.**

[17] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho and S. W. Baik, "*Convolutional Neural Networks Based Fire Detection in Surveillance Videos*," *in IEEE Access*, **Vol. 6, pp. 18174-18183, 2018.**

AUTHORS PROFILE

Dr. Julie M. David completed her Masters Degree in Computer Applications and Masters of Philosophy in Computer Science in the years 2000 and 2009 in Bharathiyar University, Coimbatore, India and in Vinayaka Missions University, Salem, India respectively. She has also completed her Doctorate in the research area of Artificial Intelligence from Cochin University of Science and Technology, Cochin, India in 2013. During 2000-2007, she was with Mahatma Gandhi University, Kottayam, India, as Lecturer in the Department of Computer Applications. Currently she is working as an Assistant Professor in the Department of Computer Applications with MES College, Aluva, Cochin, India. She has published several papers in International Journals and International and National Conference Proceedings. Her research interests include Artificial Intelligence, Data Mining, Deep Learning and Machine Learning. She is a life member of International Association of Engineers, IAENG Societies of Artificial Intelligence & Data Mining, Computer Society of India, etc. and a Reviewer of Elsevier International Journal of Knowledge Based Systems. Also, she is an Editorial Board Member of various other International Journals. She is a recipient of best paper awards in many occasions, in international events. She has coordinated various International and National Conferences.

Sarika S, born in 1997, received her BSc Degree and MSc Degree in Computer Science from Mahatma Gandhi University, Kottayam, India in 2018 and 2021 respectively. Her research interest includes Artificial Intelligence and Data Mining.

