International Journal of Engineering and Applied Physics (IJEAP)
Vol. 1, No. 2, May 2021, pp. 174~178
ISSN: 2737-8071


# Cardiovascular Ailments Prediction and Analysis Based On Deep Learning Techniques

**Riddhi Kasabe[1], Dr. Prof. Geetika Narang[2],**
[1,2]Dept. of Computer science Engineering, KJEI's Trinity College of Engineering and Research, Pune, Maharashtra.

**Article Info**

***Article history:***

Received Sep 17, 2020
Revised Feb 11, 2021
Accepted Feb 21, 2021

***Keywords:***

Heart diagnosis
Machine Learning
Data Mining
Naive Bayes
Decision trees
Classification

**ABSTRACT**

The process of data analyzing from various perspectives and combining it into useful information is called Data mining. It is used for effective prediction of heart ailment. It will be based on risk factor the heart ailments that can be defined very easily. The main objective of this project is to evaluate different classification techniques in heart diagnosis. Firstly, the heart numeric dataset is extracted and preprocessed. Then, using extraction the features that are conditioned, are found to be classified by machine learning. Compared to existing system; machine learning provides better results and efficiency. Post steps like data classification, data precision, performance criteria involving accuracy F-measure is to be calculated. Machine learning provides better results and performance of the system. The comparison measure signify that Random Forest is the best classifier that can be used for the diagnosis of heart ailment on the existing sample dataset.



***Corresponding Author:***

Dr. Prof. Geetika Narang,
HOD Computer Engineering
KJEI's Trinity College of Engineering and Research, Pune
Email: geetika.narang@gmail.com

## 1. INTRODUCTION

The main cause of deaths every year worldwide including the continent of South Africa is due to the lack of early diagnosis and predictive analysis of heart ailments. With a wealth of concealed information present, medical practitioners generate data and it is not used effectively for predictive analysis. Hence, using different data mining techniques the research converts the unused data into a dataset for shaping. The symptoms that are not taken into consideration cause many deaths yearly. There is a requirement for medical practitioners to defined heart ailment before they occur in their patients.The features that increase the probability of heart attacks are lack of physical exercises, smoking, , high cholesterol, unhealthy diet, high blood pressure, high sugar levels and detrimental use of alcohol. Cardio Vascular Disease (CVD) mainly comprise coronary heart, hypertensive heart ailment, cardio-vascular or Stroke, peripheral artery, congenital heart, inflammatory heart ailment and rheumatic heart ailment.Data mining is a technique involving knowledge discovery to encapsulate data into useful information by examining data. The current research intends to effectively predict and analyze the possibility of getting heart ailments/attack in the given patient data set. Principal goals of data mining; in practice are, Prophecies' and Descriptions. In order to locate unknown or future state values of other attributes, predictive analysis is done in data mining involving attributes or variables in the data set. Discovery of patterns that describes the data to be interpreted by humans is described

## 2. RESEARCH METHOD

The classification tree ways in the data processing. The target of this paper is to explore the investigational results of the performance of varied categorization techniques for a cardiopathy dataset. The classification tree algorithms used and experimented during this work area unit Random Forest, call Stump and LMT Tree algorithmic program.[1]

In heart diseases, accurate diagnosis is primary. However, the traditional approaches are inadequate for accurate prediction and diagnosis. In the UCI data set, 297 instances of patient records, in total, are considered of which 252 records are used for training and the remaining for testing. The results have been located to be satisfying based on the assessment [2]

Stratifying HF patients according to their risk of disease relapse and consequent hospital admission would be useful from both the clinical and economic standpoints. Identifying those HF patients at high risk of hospital admission would be useful for the clinicians since they could focus on the management of such patients to prevent potential disease relapse. From the point of view of health care planning, this information would be useful in the allocation of economic resources. However, even though the availability of a large amount of data would be a great opportunity to characterize better patients suffering from chronic disease, it has been shown that it is extremely difficult to transform such complex information in useful knowledge. Machine learning techniques (MLTs) offer a new possibility in terms of the management of this information. A growing body of literature shows MLT applications in cardiology, especially for developing prediction models using both supervised and unsupervised methods. In recent years, MLTs have been increasingly used also in the field of HF research. Those fields most frequently investigated using MLTs are the identification and classification of HF cases, prediction of HF treatment adherence, prediction of HF-related adverse events, and prediction of hospital admission/readmissions of HF patients . Prediction of hospital admission/readmission of HF patients and, in general, heart disease patients, is of great interest given the healthcare resource burden related to hospital admission/readmission. [3]

Modern healthcare systems, electronic health records, and machine learning (ML) techniques allow us to mine data to select the most significant variables (allowing for reduction in the number of variables) without compromising the performance of models used for prediction of readmission and death. Moreover, ML methods based on transformation of variables may potentially further improve the performance. Objective to use ML techniques to determine the most relevant and transform variables for the prediction of 30-day readmission or death in HF patients. We identified all Western Australian patients aged 65 years and above admitted for HF between 2003–2008 in linked administrative data. We evaluated variables associated with HF readmission or death using standard statistical and ML based selection techniques. We also tested the new variables produced by transformation of the original variables. [4]

The objective of this paper is to provide prototype using big data and data modelling techniques. It can be also used to extract patterns and relationships from database, which associated with heart disease. This system consists of two databases namely, original big dataset and another is updated one. A java-file system named HDFS used to provide a user with reliable. This system can assist the healthcare practitioners to make intelligent decisions. The automation in this system would be advantageous [five]

Authors have worked to form the system a lot of accurately predict the presence of cardiovascular disease with adscititious attributes of the disease using association rules. [6]

This paper projected processing techniques to predict diseases. It will provide the survey of current techniques to extract data from dataset and it will be useful for attention practitioners. The performance is also obtained based on time taken to develop the decision tree for the system [7]

The techniques and algorithms can be directly used on a dataset for creating some models or to draw vital conclusions, and inferences from the dataset. Common attributes used for heart disease are Age, Sex, Fasting Blood Pressure, Chest Pain type, Resting ECG(test that measures the electrical activity of the heart), Number of major vessels colored by fluoroscopy, Thrust Blood Pressure (high blood pressure), Serum Cholesterol (determine the risk for developing heart disease), Thalach (maximum heart rate achieved), ST depression (finding on an electrocardiogram, trace in the ST segment is abnormally low below the baseline), painloc (chest pain location (substernal=1, otherwise=0)), Fasting blood sugar, Exang (exercise included angina), smoke, Hypertension, Food habits, weight, height and obesity.[8]

A feature extraction method was performed using Classifier Subset Evaluator by applying a training classification data to estimate the accuracy of these subsets for all used classifiers on the HD dataset and measure the quality of the generated subsets in order to evaluate the classification performance after selecting the relevant attributes per classification algorithm [9]

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. [10]

A novel Heart ailment/attack prediction and analysis mechanism, which in the beginning learns deep features and then trains these learned features is proposed. When trained with same training samples, results of experiments show the classifier, compared to all other classifiers, and correct attributes, performs quite satisfactorily. Statistics prove that the demonstration has given better results as compared to other demonstrations and found to be significant. Predictive analysis a of heart attack comprising a low population and a high dimensional dataset is difficult due to inadequate samples to learn an accurate mapping amongst class labels and features. Currently its handling is done using feature creation and selection under data classification. Naive Bayes and Random Forest preferred, as it is able to identify the underlying.

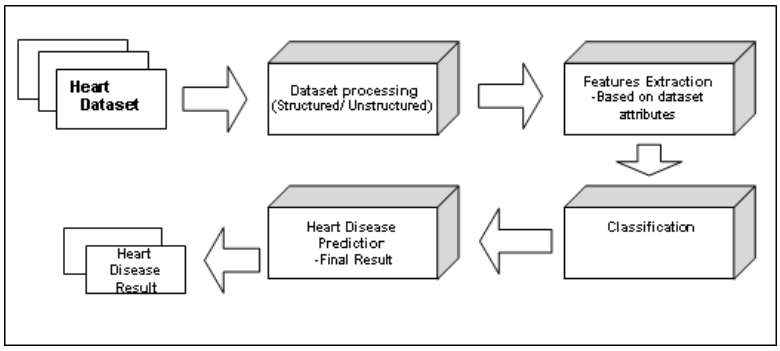


Fig. 1. Proposed System Architecture

***Algorithms***

**1.Decision Tree Algorithm**

Input:

Step 1: Upload sample dataset

Step 2: Symptoms will be the set of input attributes

Step 3: Heart ailment prediction will be the set of output attributes

Step 4: Sample will be a set of training data Function Iterative Dichotomiser will return a decision tree

1. Create root node for the tree
2. If (all inputs are positive, return leaf node positive)

If Else (if all inputs are negative, return leaf node negative) Else (Some inputs can be positive and some inputs can be negative, then check condition (Positive ¿ negative ; positive ¡ negative), thenReturn result)

3. The entropy of current state H(S) is calculated

4.Calculate the entropy with respect to the attribute 'X' denoted by H(S, X), for every attribute.

5.The attribute with maximum value of IG(S, X) is to be selected

6.The attribute that offers highest value from the set of attributes should be removed

7.Repeat the steps until the system exhausts all attributes or the decision tree has only leaf nodes.

Output:

Retrieved value will be the value of dataset.

## 3. RESULTS AND DISCUSSIONS

We compared the proposed heart ailments prediction accu- racy on number of samples and results are displayed graphically. On analysis of the following graph and table, the heart ailment prediction accuracy result, based on decision tree classification technique is displayed.

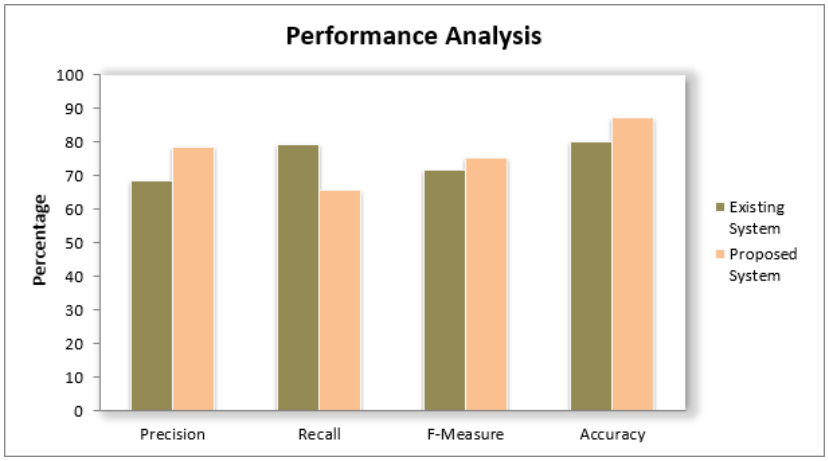


Figure 2. Analysis Graph

TABLE I ANALYSIS TABLE

| | Existing System | Proposed System |
|---|---|---|
| Precision | 68.45 | 78.70 |
| Recall | 79.44 | 65.64 |
| F-Measure | 72.11 | 74.31 |
| Accuracy | 80.29 | 87.26 |

## 4. CONCLUSION

The experiment organization is based on the dataset of Heart Ailment. This involves utilization of neural network modes that are single and multilayered. Heart Ailment dataset is considered and analysis is performed on the sample dataset to predict the intensity of the ailment. A convolution neural network approach is used to predict the intensity of the diagnosed ailment. Preprocessed data from the dataset is utilized to make it suitable for classification. The proposal of convolution neural network approach is to generate efficient classification rules. Convolution's technique is used to train the neutral network in order to perform classification on available medical data. The technique of Convolution neural network is a multilayer perceptron that is the special design for identification of 2D image information. They usually have more layers: input layer, convolution layer, sample layer and output layer. Also, in deep network, architecture samples the convolution layers and sample layers can be multiple.

## ACKNOWLEDGEMENTS

I am profoundly grateful to Dr. Mrs. Geetika Narang for her expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion. At last, I must express my sincere heartfelt gratitude to all the staff members of Department who helped me directly or indirectly during this course of work.

## REFERENCES

[1] Vijiyarani, S., and S. Sudha, "An efficient classification tree technique for heart disease prediction", International Conference on Research Trends in Computer Technologies (ICRTCT-2013) Proceedings published in International Journal of Computer Applications (IJCA)(0975–8887). Vol. 201, 2013.
[2] Mohan S, Srivastava CTAG. Effective Heart Disease Prediction using Hybrid Machine Learning Techniques. IEEE Access. 2016;4:1–14. https://doi.org/10.1109/ACCESS.2019.2923707.
[3] Comparison of Machine Learning Techniques for Prediction of Hospitalization in Heart Failure Patients Giulia Lorenzoni,1 Stefano Santo Sabato,2 Corrado Lanera,1 Daniele Bottigliengo,1 Clara Minto,1 Honoria Ocagli,1 Paola De Paolis,3 Dario Gregori,1,* Sabino Iliceto,4 and Franco Pisanò3
[4] Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death Saqib E Awan, Mohammed Bennamoun, Ferdous Sohel, Frank M Sanfilippo, Benjamin J Chow, Girish Dwivedi
[5] Prajakta Ghadge, Vrushali Girme, Kajal Kokane, Prajakta Deshmukh, "Intelligent Heart Disease Prediction System using Big Data", International Journal of Recent Research in Mathematics Computer Science and Information Technology, vol.2, October 2015 - March 2016, pp.73-77.
[6] Asmi, Shabana P., and S. Justin Samuel. "An analysis and accuracy prediction of heart disease with association rule and other
[7] Sharan Monica.L, Sathees Kumar.B, "Analysis of CardioVasular Disease Prediction using Data Mining Techniques", International Journal of Modern Computer Science, vol.4, 1 February 2016, pp.55-58.
[8] Muthuvel, Marimuthu & Abinaya, M & Hariesh, K & Madhankumar, K & Pavithra, V. (2018). A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach. International Journal of Computer Applications. 181. 975-8887. 10.5120/ijca2018917863
[9] Prediction of heart disease and classifiers' sensitivity analysis Khaled Mohamad Almustafa BMC Bioinformatics volume 21, Article number: 278 (2020)
[10] Performance Analysis of Convolutional Network System for Heart Disease Prediction Julie M. David1*Sarika S. DOI: https://doi.org/10.26438/ijcse/v9i2.1722 | Available online at: www.ijcseonline.org

## BIOGRAPHIES OF AUTHORS


| | |
|---|---|
|  | Riddhi Kasabe<br>PG Student<br>KJEI's Trinity College of Engineering and Research, Pune |
|  | Dr. Prof Geetika Narang<br>Hod Computer Engineering<br>KJEI's Trinity College of Engineering and Research, Pune |