International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 02 | Feb 2020 www.irjet.net p-ISSN: 2395-0072

Heart Failure Risk Prediction using Trained Electronic Health Record

T. Jenifer[1], M.I. Peter shiyam[2]

[1]*PG Student, Computer Science & Engg., DMI Engineering College, Tamilnadu, India*
[2]*Assistant Professor, Computer Science & Engg., DMI Engineering College, Tamilnadu, India*

Abstract - *Heart related issues are terribly high and it can't be simply predict earlier to avoid complications. Heart failure is a modern issue in cardiology. This paper's aim is to use Electronic Health Record to determine the risk of patients having a heart failure. EHR mainly contains information about the patient's lab and test results, medical history and Administration details. We get huge data from electronic health record. We can easily predict the threat of a heart failure using these results. This paper propose K-dimensional tree classifier to classify data and predicting the risk of heart failure. This paper's primary objective is to quantify the risk and to reduce the classification interruption.*

***Key Words*: Electronic Health Record, Heart Failure, K-dimensional Nearest Neighbour.**

1. INTRODUCTION

Heart failure may be a significant issue in today's word, and it cannot realize simply in early stage. The most reasons for heart failure Smoking, unhealthy diet, High pressure, polygenic disorder and then on. To rectify the heart failure risk we need to analyze the patient's medical data. The most risk factors heart failure are High pressure, heart failure, diabetes, Some polygenic disorder medications, certain medications, sleep disorder, etc. within the previous system they use less quantity of options. Using the less quantity of options we can't predict the risk accurately. And that they use WPCA methodology for classification and feature extraction. WPCA would like standardization otherwise principal element analysis won't be able to realize the best Principal parts. Using this methodology sometimes data's may be loses. To beat these we propose k-dimensional tree for classification. This is often a One of the kind of KNN formula. KNN is loosely classified into two varieties 1) structure less NN techniques 2) structure primarily based NN techniques. In structure less NN techniques complete information was classified through training and sample information is reviewed. Structure primarily based NN techniques are supported structures of knowledge. Here we tend to use k-d tree classifier. K-d tree is a structure primarily based technique. Using these formulae we can simply check the probability of heart failure.

And here we are using more risk factors for risk prediction. Chest pain, fasting blood sugar, cholesterol, fetal bovine serum, resting Electro cardio graphic tests, maximum heart rate reached, angina induced exercise, old peak, number of major vessels, resting heart rate are the risk factors used here. Those are the critical heart failure characteristics. Table-1 shows heart failure deficiency characteristics. We use k-d tree here.

Table -1: Heart Failure Attributes

No	Attributes	Explanation
1.	Age	Patient Age.
2.	sex	Male=1 Women=0
3.	cp	A pain type in the chest.
4.	Trestbps	Residual blood pressure recorded in mm Hg, during hospital admission.
5.	Chol	The patient reported cholesterol in mg / dl.
6.	Fbd	Fasting glucose greater than one hundred twenty 1=true 0=false
7.	Restecg	Resting graphical record result
8.	Thalach	Maximum pulse rate achieved
9.	Exang	Exercising caused angina.
10.	Oldpeak	ST depression evoked by exercise relative to rest.
11.	Slope	Measure of slope for peak exercise.
12.	Ca	No.of major vessels.
13.	Thal	Type of fault.
14.	Target	The foretold attribute

2. LITERATURE SURVEY

In [1] Abhishek Taneja "Heart Disease Prediction System Using Data Mining Techniques" The intention of the scientific research was to establish a predictive model for heart failure identification which use data mining strategies from the Transthoracic Echocardiography Study data sets that could increase the efficiency of echocardiography treatment of cardiac disease. Based on the preprocessed data set, the designs were designed using WEKA 3.6.4 machine learning software, using three various supervised machine learning algorithms J48 Classifier, Naive Bayes and Multilayer Perception.

In[2] Bhatla N, Jyoti K "An Analysis of Heart Disease Prediction using Different Data Mining Techniques" Established a desktop-aided heart disease diagnosis system that helps the physician as a method for identifying heart disease. The study shows that neural network with offline planning is good for early-stage disease diagnosis and strong performance can be achieved by pre-processed and structured set of data.

In[3] B. Jin, Senior Member, C. Che, Z. Liu, Shulong Zhang, Xiaomeng Yin and X.P. Wei, "Predicting the risk of Heart Failure with EHR sequential data modelling" This suggests an important and reliable system for predicting heart failure. The Biggest contribution of the whole paper is to find out heart problem using a neural network. Using the fundamental principles of a long-term network memory model, they used single-hot encoding and word vectors to image the diagnostic events and predicted heart failure events.

In[4] Chen A.H, Huang S.Y, Hong P.S, Cheng C.H, E J Lin, "HDPS: Heart Disease Prediction System" Build a predictive algorithm for heart disease that can help healthcare professionals predict the status of heart disease based on patient clinical data. The method involves three key steps involving the collection of 13 essential clinical characteristics in the first phase. Phase two was to build an ANN algorithm based on features to identify heart disease and third phase to create a user friendly heart disease prediction method.

In[5] Jabbar M.A., Deekshatulu B. L.,Chandra P., "Heart Disease Prediction using Lazy Associative Classification" Initiated a lazy associative categorisation for prediction of heart disease in Andhra Pradesh, and presented some scientific findings using 7 UCI Repository data sets. Research scientist employed centric intelligence attribute test PCA to generate rule of class association that will be used to calculate the occurrence of heart disease. Scientists are developing the framework for the population of Andhra Pradesh because this state is at risk for most deaths from heart disease.

In [6] Jyoti Soni et.al. "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction" Naive Bayes as well as a classification system to identify the occurrence of heart disease in clinicians. Clustering: Clustering is the process through which related components are clustered together. For experiments the WEKA 3.6.0 programming was used. 909 Data Set documents, with 13 characteristics.

In [7] KarayÕlan T., KÕlÕç O., "Prediction of Heart Disease Using Neural Network" Studied complex software methods that help detect various heart diseases. Few of the mechanisms used are Hidden Markov Models, Support Vector Machine, Computational Intelligent Classifier, and Data Mining etc. Because the treatment of cardiac disorders is very costly and inexpensive to any normal individual, these advanced technology styles are designed to overcome this problem. Such methods are also useful for early phase predictions. Through making slight improvements in the daily routine it eliminates any other potential sufferings. Accordingly, the author concludes that the expected solution has several benefits and is very useful.

In[8] V. Manikandan and S. Latha, "Predicting the Analysis of Heart Disease Symptoms Using Medical Data Mining Methods" Suggests that association rule mining is being used to retrieve the connections established for the subject. Data classification was dependent on MAFIA algorithms which resulted in greater ease. The outcomes were analyzed through entropy-based cross-validation and partitioning methodology, and the outcomes were compared. MAFIA (Maximal Frequent Item Set Algorithm) used a dataset of 19 attributes and the purpose of the scientific research was to have extremely precise retrieval measurements of higher precision levels.

In[9] Nidhi Bhatla, Kiran Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques". In their research work they relied on using different algorithms and combinations of multiple target attributes to estimate heart attacks reliably using data mining. Decision Tree has outperformed 99.62 percent precision by using 15 parameters. Upon introducing the genetic algorithm to both the actual data scale, the efficiency of the Decision Tree and Bayesian Classification also improves further in order to acquire the correct range of parameters required for detection of heart disease.

In[10] Polat K., Sahan S., Gunes S., "Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn based weighting preprocessing" Implemented repository knowledge which has been fed up into multiple classifiers such as KNN, naïve bayes, SVM and decision trees, each undertaking specific tasks. Similar methods are used in early diagnosis of heart disease. To achieve the right and optimal diagnosis for heart-related diseases, it requires very successful results. This approach has certain benefits, such as being able to predict the diseases reliably at their very early stages, and being handled on time correctly and efficiently.

3. SYSTEM ANALYSIS

3.1 Existing System

Weighted Principle Component Analysis (WPCA) procedure is being used for classification within the previous work. A lot of tools are used to determine the risk of heart failure. This requires an orthogonal transition to reduce a variety of possible measures of associated substances into a bunch of guidelines of linearly uncorrelated quantities called primary elements. The existing system has some disadvantages:

- Existing studies do not include more features of heart disease.
- If we do not pick the no. of principle components with care it loss information.
- Information loss is high.

3.2 Proposed System

To overcome the existing system, the proposed system uses K-d tree nearest neighbor (KDNN) classifier for predicting the risk of heart failure. KDNN is a type of nearest neighbor algorithm. The k-dimensional trees divide the trained data into two parts, to find out the corresponding result. The proposed system has following advantages:

- More features are being used to estimate the risk.
- Enables the algorithm to adapt quickly to input changes.
- Easy to implement and follow.

4. SYSTEM DESIGN & IMPLEMENTATION

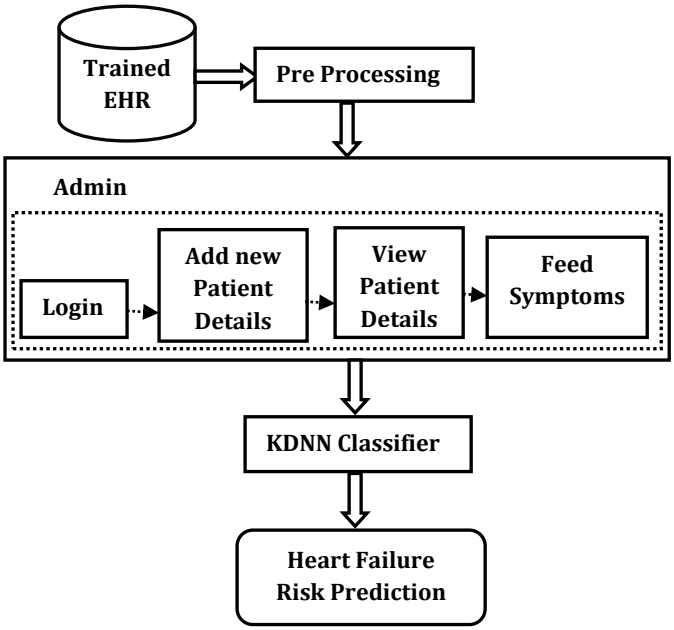

Fig -1: System Architecture

4.1 Module

- Trained EHR
- Pre-processing
- Admin registration module
- KDNN classifier
- Risk Prediction

4.2 Trained EHR

The Trained EHR includes the medical history, diagnosis and reports of the patient's laboratory.

4.3 Pre-processing

Preprocessing refers to the data transformation that was applied before the algorithm was feed. Convert the relevant data in to the clean set of data.

4.4 Admin registration module

With the correct user_id and password Admin has to login into homepage. Admin performs the following: adding new patients detail, view the patient details and feed the symptoms.

- **Add new patients**: Adding new patient details.
- **View patient:** Viewing the patient details.
- **Feed symptoms**: Feeding the symptoms of a patient.

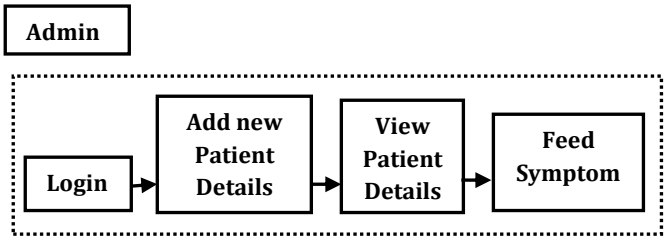

Fig -2: Admin login

4.5 KDNN classifier

The k-dimensional trees partition the EHR data set into two parts, the right portion and the left portion. It analyses the left or right side according to data on the issue. Once the terminal node has been reached, records in the terminal node are checked to find the nearest data node to query information.

If (searched data in node_left)

Return "YES"

If (searched data in node_right)

Return "YES"

Else:

Return "NO"

If system return "YES", then visualization of the result will show.

4.6 Risk Prediction

The final step is to predict risks. In this step we can verify whether or not the risk is present. If the risk is present it

will give a detected message of risk otherwise it will give risk not detected.

5. RESULTS

Fig -3: Login page

Fig -4: Registration page

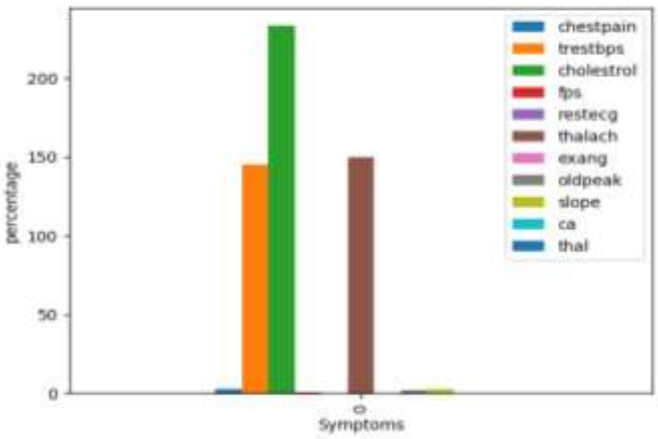

Fig -5: Risk Prediction

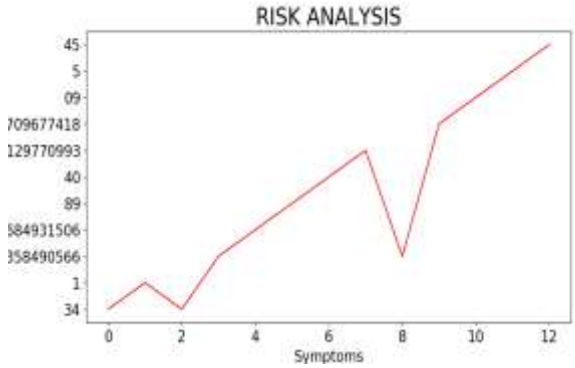

Fig -6: Risk analysis graph

6. CONCLUSIONS & FUTURE WORK

The main aim is to predict heart failure risk using classification technique. We presented K-dimensional tree nearest neighbor (KDNN) classifier for classification. These method divides the trained EHR into two part right node and left node. Left or right side of tree is searched according to input data. The result shows that the K-dimensional nearest neighbor approach can reliably predict the risk of heart failure.

Based on the current work there are several prospective paths worth exploring in the future. In future we can predict the level of risk i.e. high or low and check each features of heart failure patient. It analyze each and every data which provides the output as the heart failure risk whether it is high or low. In future the training dataset is used for prediction and also predict the accuracy level.

REFERENCES

[1] Abhishek Taneja, Heart Disease Prediction System Using Data Mining Techniques; Oriental Journal of computer science & Technology ISSN: 0974-6471 December2013.

[2] Bhatla N., Jyoti K., "An Analysis of Heart Disease Prediction using Different Data Mining Techniques" International Journal of Engineering Research & Technology, Volume 1, Issue 8, 2012.

[3] B. Jin, Senior Member, IEEE , C. Che, Z. Liu, Shulong Zhang, Xiaomeng Yin and X.P. Wei, "Predicting the risk of Heart Failure with EHR sequential data modelling", DOI 10.1109/ACCESS.2017.2789324, IEEE Access

[4] Chen A.H, Huang S.Y, Hong P.S, Cheng C.H, E J Lin, "HDPS: Heart Disease Prediction System", IEEE, 2011.

[5] Jabbar M.A., Deekshatulu B. L.,Chandra P., "Heart Disease Prediction using Lazy Associative Classification", IEEE, 2013.

[6] Jyoti Soni et.al. Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction; International Journal of Computer Applications (0975 – 8887) Volume 17– No.8, March 2011.

[7] KarayÕlan T., KÕlÕç O., "Prediction of Heart Disease Using Neural Network", IEEE, 2017.

[8] V. Manikandan and S. Latha, "Predicting the Analysis of Heart Disease Symptoms Using Medical Data Mining Methods "International Journal of Advanced Computer Theory and Engineering", Vol. 2, Issue. 2, 2013.

[9] Nidhi Bhatla, Kiran Jyoti, " An Analysis of Heart Disease Prediction using Different Data Mining Techniques" International Journal of Engineering and Technology Vol.1 issue 8 2012.

[10] Polat K., Sahan S., Gunes S., "Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing", ScienceDirect, 2007.

[11] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

AUTHORS

T.JENIFER is currently pursuing her M.E degree in Department of Computer Science & Engineering at DMI Engineering College, Aralvoimozhi, and Tamilnadu, India.

M.I.PETER SHIYAM, Currently working as assistant professor, Dept. of Computer Science & Engg. DMI Engineering College, Aralvoimozhi, Tamilnadu, India.