RESEARCH ARTICLE   OPEN ACCESS

Prediction of Heart Disease Using Enhanced Association Rule Based Algorithm

Karandeep Kaur*, Ms. Poonamdeep Kaur**, Ms. Lovepreet Kaur***
*(Student (Computer Science & Engineering), Guru Nanak Dev Engineering College, Ludhiana)
** (Assistant Professor (Computer Science & Engineering), Guru Nanak Dev Engineering College, Ludhiana)
*** (Assistant Professor (Information Technology), Guru Nanak Dev Engineering College, Ludhiana)

--✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳--------------------------------

Abstract:

The risk of coronary illness is increasing at a fast pace and it's been scowling at us for a considerable length of time, making us doubt each subtle element of our confounded way of life decisions, eating regimen and level of physical movement. It's been a main executioner in the West and has now forcefully advanced toward India. As indicated by government information, the pervasiveness of heart disappointment in India because of coronary illness, hypertension, corpulence, diabetes and rheumatic coronary illness ranges from anyplace between 1.3 to 4.6 million, with a yearly frequency of 491,600 to 1.8 million. But because of impreciseness of the diagnosis tools less than 68 per cent of heart diagnosis yield correct results. To make the odds go higher, this research presents a novel algorithm whose key idea is germinated from classical association rule mining.

Keywords — **Association Rule Mining, Heart Disease, UCI Machine Learning.**

--✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳✳--------------------------------

I.   INTRODUCTION

Mechanical progressions and social insurance mindfulness have driven towards the improvement of tremendous number of human services offices and doctor's facilities. On the other hand, given the high caliber of medicinal services administrations requiring little to no effort is turning into a testing issues inside the developing nations across the globe albeit, numerous nations have made some firm strides towards the guaranteeing that human services administrations are given to everybody. Therapeutic information mining has an incredible capacity for the investigation of the shrouded examples in the current datasets of the medicinal area. Such examples could be used with the end goal of clinical conclusion. A noteworthy test that this Industry's appearances is giving quality administrations at reasonable costs. Quality administration suggests diagnosing an understanding's condition viably, giving fitting medications and observing those medicines all the time. Uncalled for clinical choices may prompt sad results. Alongside diagnosing conditions and giving suitable medications, healing centres should likewise minimize the expense of clinical tests. A successful method for accomplishing these outcomes is by utilizing proper PC - based data and/or choice emotionally supportive networks. Most doctors' facilities today utilize some kind of data frameworks to deal with their medicinal services or patient information. These frameworks create colossal measures of information as numbers, outlines, writings pictures and so forth. Gigantic measure of information is put away on consistent premise however sadly such information is infrequently used to bolster clinical choice making. There is a substantial number of shrouded data in these information that is to a great extent undiscovered.
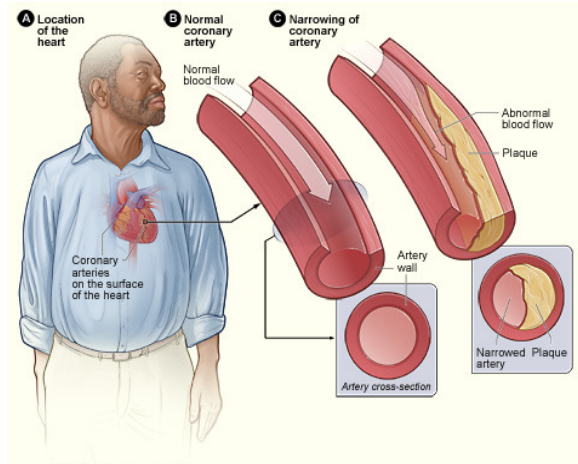
Fig.1: Coronary Heart Disease

Information mining obliges an accumulation of information in a sorted out shape, the information gathered was incorporated in the arrangement of a clinic data framework. The innovation of information mining gives the clients a client arranged methodology towards concealed and novel examples in information. Proficient and powerful robotized coronary illness framework fit for foreseeing coronary illness could turn out to be tremendously useful for the social insurance segment. This research study endeavours to present a point by point investigation of diverse information of Enhanced Association Rule Learning for the expectation of coronary illness using the information comprising of a person's medical history. These mechanizing frameworks will assume a noteworthy part in lessening the general number of tests that a patient needs to take concerning coronary illness.

II. METHODOLOGY

This paper formulates a novel algorithm based on association rule learning to precisely predict the result of a coronary illness examination and the act of this novel calculation could turn out to be useful for the medicinal specialists and experts for precisely anticipating the coronary illness.

A. Inputs

13 attributes taken from 270 different observations is the input to the proposed system. The type of each attribute is given below:

Attribute	Type
Age	Quantitative
Sex	Categorical
Chest Pain Type	Categorical
Resting Blood Pressure	Quantitative
Serum Cholesterol	Quantitative
Fasting Blood Sugar	Categorical
Resting Electrocardiographic results	Categorical
Maximum Heart Rate Achieved	Quantitative
Exercise Induced Angina	Categorical
Oldpeak	Quantitative
ST segment	Categorical
Coloured Vessels	Categorical
Thal	Categorical
Result	Categorical

B. Algorithm

The algorithm of the proposed methodology is given below:

Step 1: Scan x=1 to n
Step 2: Set Min(Support) = μ
Step 3: if (Support < μ) goto step 6
Step 4: S = Hypothesis / (Total number of rows)
Step 5: C = Conclusion / (Total number of rows)
Step 6: Stop
Step 7: Compare logarithm of odds of happening of heart disease to linear function
Step 8: Compute Probability
Step 9: Stop

C. Explanation

The dataset for this research is taken from UCI machine learning repository, and the dataset name is Statlog for heart disease prediction.

1) **Attributes:** The dataset contained 13 different attributes from which the result has to be predicted. There are exactly 270 cell values associated with each attributes. These

attributes are scanned at first and then minimum support is estimated by observing the dataset.

2) Support greater than Minimum Support: If the support is greater than minimum support then confidence and support is calculated for the given dataset and the result is predicted.

3) Support less than Minimum Support: In this case, the Joint estimated probability of the given data is computed. And based on those computed probability the results are again predicted.
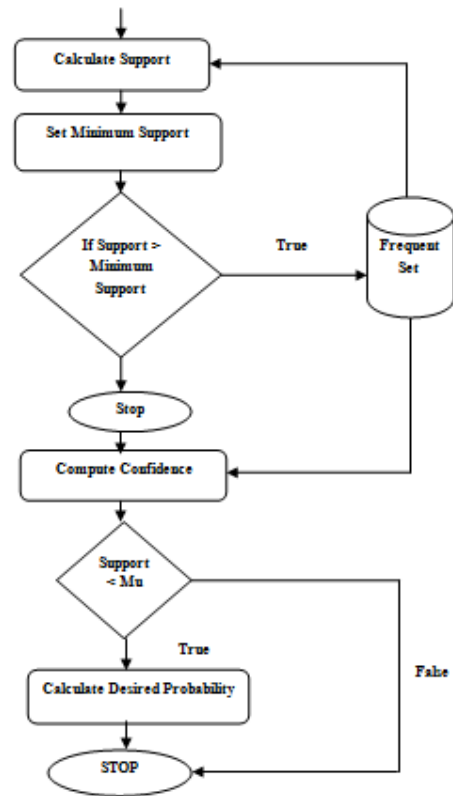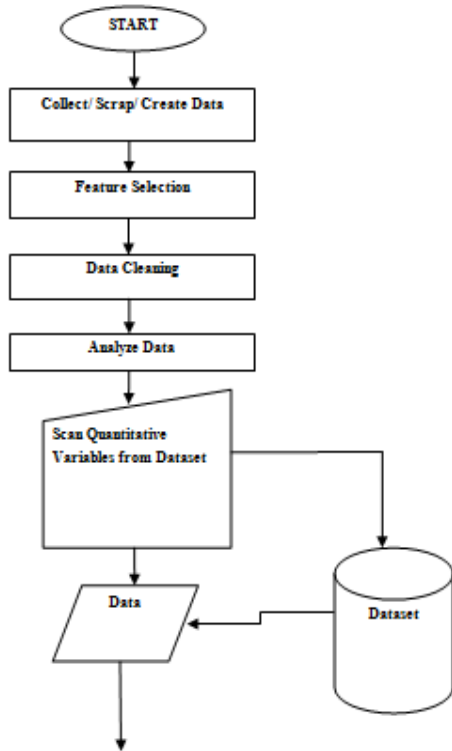
D. Flow Chart

Fig. 2: Flow Chart of Proposed Model

III.   RESULTS

To test the accuracy of the system, 500 different test cases were randomly chosen, out 500 test samples 50 test samples were taken from the original test cases themselves.

To contrast the efficiency of the proposed system, 5 best algorithms were taken namely, Associative Classification and Hybrid Feature Subset Selection, Naïve Bayes, C4.5, k-Nearest   Neighbor, and Artificial Neural Network. And the proposed methodology turns out to be more efficient than others. The comparison is shown in the table below:

Table 1: Comparison Table

ACHFSS	Naïve Bayes	C 4.5	KNN	ANN	Proposed Method
95	65	77.5	97.4	94	97.6

The comparison of the proposed system is based on their accuracy for 500 data items taken from different data sources. The formula used for the evaluation for the accuracy is given by:

**Accuracy= (Number of correct prediction+
Number of incorrect prediction)
/ (Total number of cases)**

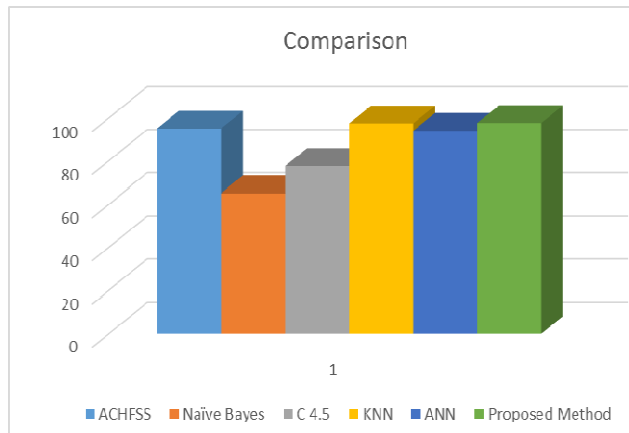The bar chart representation of the comparison is shown the figure below:

Fig. 3: Comparison

IV. CONCLUSION

The target work is to anticipate the risk score of coronary illness from the data with high exactness .We utilized upgraded affiliated principle to focus the presence of affliction in a man. The significant issue with affiliated order is that it works better for clear cut antecedent yet for our situation a large portion of the forerunner present in the data source was quantitative in nature. Furthermore, Associative rule mining, works best for the cases were no example or test outcome are available on the grounds that associative rule mining is an unsupervised learning calculation. What's more, it is constantly shrewd to pick directed learning calculation to foresee scores if some specimen result set are accessible. Yet, administered learning calculation accompanies a punishing expense, and that it is exceptionally time wasteful and

computational costly and if legitimate seeds and weights are not balanced they may not even unite, consequently they require expansive measure of information set and mount focuses ought to be picked by a specialist to make the procedure meet appropriately. Accordingly in this theory associative rule is improved by changing it into a force managed learning calculation. Thusly, now it yields preferred result over some other calculation exhibit in the writing without losing its sharp precision and expedient conduct. The proposed strategy can be utilized as a part of different other disciplinary work to acquire precision with no speed punishment.

V. FUTURE SCOPE

As the Healthcare area is dynamic and this issue is a test to the information mining. It is additionally a driving inspiration to the information mining applications in human services. This dynamism offers approach to new skylines and more information mining applications will be utilized to find new examples and affiliations. In the perspective of the subjects analyzed in this study, future information mining studies appear to happen and not restricted but rather in impressive weight, in circulated information mining applications and content mining calculations. With the assistance of information mining calculations, the grouping execution increments. This can be further improved and extended with more expectation calculation for significant life debilitating infections. In this thesis, we have improved the technique for successful heart assault forecast framework utilizing information mining and the enhanced associative rule mining which gives more precise result. It is desktop application where running on stand alone in desktop framework. In future work, we mean to apply on superior customer server or parallel architectures and optical neural system as a classifier model. Utilizing information mining strategies to help human services experts in diagnosing and giving suitable medications to coronary illness and Continuous information can likewise be utilized rather than simply unmitigated information. The customer server application introduced that demand and gets data over the

system so it would simple to get to this application to all clients.

ACKNOWLEDGMENT

REFERENCES

[1]   V. V. R. Bangaru Veera Balaji, "Improved Classification Based Association Rule Mining," *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 2, p. 11, May 2013.

[2]   T. K. and P. S. Chobe, "An Overview of Association Rule Mining Algorithms," *International Journal of Computer Science and Technologies,* vol. 5, p. 5, July 2014.

[3]   D. J. Hand, "Principles of data mining," pp. 621-622.

[4]   N. L. Komal Hadvani, "A Review On Early Heart Diseases Prediction Using Data Mining Techniques," *International Journal of Innovative Research in Technology,* vol. 1, p. 6, August 2014.

[5]   B. Liu, *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, 2011.

[6]   V. S. Nikita Jain "Data Mining Techniques: A Survey Paper," *International Journal of Research in Engineering and Technology,* vol. 2, p. 5, November 2013.

[7]   R. C. and V. Seenivasagam, "Review Of Heart Disease Prediction System Using Data Mining And Hybrid Intelligent Techniques," *ICTACT Journal On Soft Computing,* vol. 3, p. 6, July 2013.