



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Heart Attack Prediction System Using Data Mining

Prof. Thakur Ritesh, Bendre Gauri, Darekar Ashwini, Shinde Priyanka

Department of Computer Engineering, Savitribai Phule Pune University, Institute of Knowledge College of
Engineering, Pimpale Jagtap, Pune, India

ABSTRACT: The Heart Attack Prediction System is a web based application intended for medical field. The health care environment is found to be rich in information, but poor in extracting knowledge from the information. Data mining tools traditionally taken much time consuming to resolve. The huge amounts of data generated for prediction of heart disease are too complex and voluminous to be processed and analyzed by traditional methods. The proposed system is intended to develop an Intelligent System using data mining technique such as naïve bayes algorithm and implemented as Java application. The test results are used to compute to compute prediction result.

KEYWORDS: data mining, classification, entropy, gain, health informatics, Naïve Bayesian

I. INTRODUCTION

Coronary illness is the greatest reason for death these days. Blood weight, cholesterol, beat rate are the significant explanation behind the coronary illness. Some non-modifiable components are additionally there. Such as smoking, drinking likewise explanation behind coronary illness. The heart is a working arrangement of our human body. On the off chance that the capacity of heart is not done legitimately implies, it will influence other human body part moreover. Some danger variables of coronary illness are Family history, High circulatory strain, Cholesterol, Age, Poor eating routine, Smoking. At the point when veins are overstretched, the danger level of the veins are expanded. This prompts the blood weight. Pulse is ordinarily measured as far as systolic and diastolic. Systolic demonstrates the weight in the corridors when the heart muscle contracts and diastolic shows the weight in the corridors when the heart muscle is in resting state. The level of lipids or fats expanded in the blood are causes the coronary illness. The lipids are in the veins subsequently the corridors get to be restricted and blood stream is additionally turned out to be moderate. Age is the non-modifiable danger component which additionally an explanation behind coronary illness. Smoking is the purpose behind 40% of the passing of heart maladies. Since it constrains the oxygen level in the blood at that point it harm and fix the veins. Different information mining procedures, for example, Naïve Bayes, KNN calculation, Choice tree, Neural Network are utilized to foresee the danger of coronary illness [1]. The KNN calculation utilizes the K client characterized quality to discover the estimations of the elements of coronary illness. Choice tree calculation is utilized to give the grouped report to the coronary illness. The Naïve Bayes strategy is utilized to anticipate the coronary illness through likelihood. The Neural Network gives the minimized blunder of the expectation of coronary illness. In this previously mentioned methods the patient records are ordered and anticipated consistently. The patient action is checked consistently, if there is any progressions happen, then the hazard level of illness is educated to the patient and specialist. The specialists can foresee heart illnesses at a prior stage in view of machine learning calculations and with the assistance of PC innovation.

A. DATA MINING :

Data mining is an interdisciplinary subfield of software engineering. It is the computational procedure of finding examples in vast information sets including techniques at the crossing point of counterfeit consciousness, machine learning, measurements, and database frameworks. The general objective of the information mining procedure is to concentrate data from an information set and change it into a reasonable structure for further utilize. Beside the crude examination step, it includes database and information administration angles, information pre-processing, model and derivation contemplations, fascinating measurements, many-sided quality contemplations, post-preparing of found



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

structures, representation, and internet redesigning. Data mining is the examination venture of the "learning revelation in databases" process.

For instance, one Midwest basic supply chain utilized the information mining limit of Oracle programming to investigate nearby purchasing designs. They found that when men purchased diapers on Thursdays and Saturdays, they likewise tended to purchase brew. Further investigation demonstrated that these customers regularly did their week after week shopping for food on Saturdays. On Thursdays, nonetheless, they just purchased a couple of things. The retailer presumed that they acquired the brew to have it accessible for the forthcoming weekend. The basic supply chain could utilize this newfound data in different approaches to expand income. For instance, they could draw the lager show nearer to the diaper show. What's more, they could ensure brew and diapers were sold at the maximum on Thursdays.

B. TECHNIQUES USED IN DATA MINING :

While expansive scale data innovation has been advancing separate exchange and scientific frameworks, information mining gives the connection between the two. Information mining programming breaks down connections and examples in put away exchange information taking into account open-finished client questions. A few sorts of logical programming are accessible: factual, machine learning, and neural systems. For the most part, any of four sorts of connections are looked for:

1. Classes : Put away information is utilized to find information in foreordained gatherings. For instance, an eatery network could mine client buy information to decide when clients visit and what they normally arrange. This data could be utilized to build movement by having day by day specials.
2. Clusters: Information things are gathered by connections or customer inclinations. For instance, information can be mined to recognize market fragments or shopper affinities.
3. Associations: Information can be mined to distinguish affiliations. The lager diaper case is a case of affiliated mining.
4. Sequential patterns: Information is mined to foresee conduct examples and patterns. For instance, an open air gear retailer could foresee the probability of a knapsack being obtained taking into account a customer's buy of resting packs and trekking shoes.

II. LITERATURE SURVEY

A. ELECTRONIC RECORDING SYSTEM HEART DISEASE PREDICTION SYSTEM

The Heart Disease Prediction System (HDPS) can discover and extract hidden knowledge associated with diseases from a heart disease database. It can answer complex queries for diagnosing disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs. To enhance visualization and ease of interpretation, it displays the results in tabular forms. The system uses various data mining techniques. In our system this drawback will get overcome.

B. USING DATA MINING TECHNIQUES IN HEART DISEASE DIAGNOSIS AND TREATMENT

This paper identifies gaps in the research on heart disease diagnosis and treatment and proposes a model to systematically close those gaps to discover if applying data mining techniques to heart disease treatment data can provide as reliable performance as that achieved in diagnosing heart disease patients.

III. EXISTING SYSTEM

Choice Tree is a standout amongst the most mainstream order calculations in current use in Data Mining. Choice Tree incorporates different sorts of calculations, for example, ID3, C4.5, C5, J48 and CART. In this paper, ID3 calculation is used in light of the fact that they are most appropriate for coronary illness dataset. The essential thought of Iterative Dichotomiser 3 or ID3 calculation is to develop the choice tree by utilizing a top-down, voracious pursuit through the offered sets to test every characteristic at each tree hub. The fundamental ideas utilized as a part of ID3 calculation are Information Gain and Entropy to choose the property that is most helpful for grouping the given sets.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

But it is very time consuming and less accurate. Drawback of existing system is that, it is not able to detect heart attack in earlier stage.

Existing Algorithm with its Accuracy.

CART=83.49%

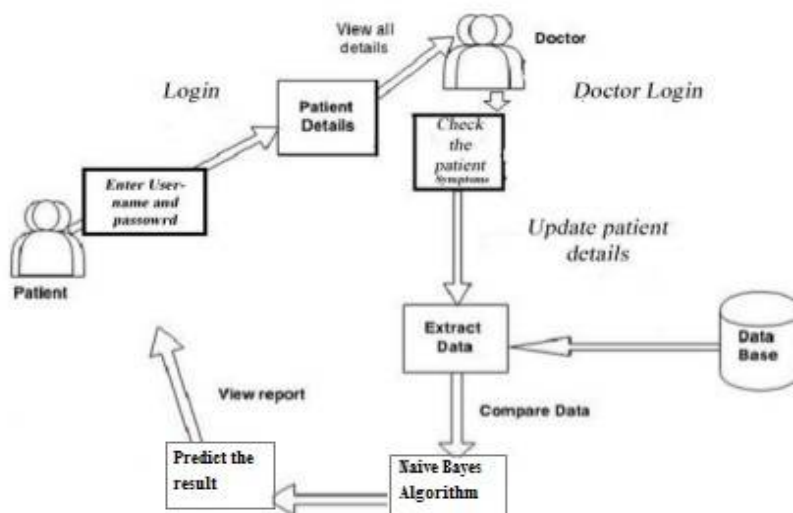
ID3=72.93%

Decision Table =82.50%

IV. PROPOSED SYSTEM

We suggest that applying data mining methods in recognizing appropriate medications for heart patients is productive and needs encourage examination. To assess if applying data mining procedures to heart attack detection can give as dependable execution as accomplished in heart attack conclusion.

In this system we are using Naive Bayes Theorem. Naive Bayes is a straight forward procedure for developing classifiers like models that dole out class marks to issue occurrences, spoke to as vectors of highlight qualities, where the class names are drawn from some limited set. It is not a solitary calculation for preparing such classifiers, but rather a group of calculations in view of a typical standard.



8

C. Work Flow :

1. Patient first do their registration and after finished his or her test that person update test result.
2. This result goes in database. In database we already have existing report of previous patients. so here naive bayes algorithm is apply. This algorithm select one base line. According to that base line system will give result. this result is shown to the doctor.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

V. APPLICATIONS

1. In medical and health care areas.
2. Disease diagnosis is one of the applications where data mining tools are proving successful results.
3. Researchers have long been concerned with applying statistical and data mining tools to improve data analysis on large data sets.
4. Treatment records of millions of patients can be stored and computerized and data mining techniques may help in answering several important and critical questions related to health care.

VI. FUTURE WORK

Here this system is introduced for Heart Attack Prediction but in future this system will implemented for all type of disease.

VII. CONCLUSION

In this paper newly Heart Attack Prediction system based on naive bayes algorithm is introduced. For this Data mining concept is used, so this paper gives a quick and basic insightful of various expectation models in data mining and finds most prominent model for further work.

This system is convenient ,effective and gives good prediction of disease to patient. This work can be improved by expanding the quantity of traits for the current arrangement of our past work.

VIII. ACKNOWLEDGEMENT

We are working on this project under the guidance of Prof. Thakur Ritesh, Associate Professor at Institute Of Knowledge College Of Engineering, PUNE.

REFERENCES

- [1] Ranganatha S., Pooja Raj H.R., Anusha C., Vinay S.K., "MEDICAL DATA MINING AND ANALYSIS FOR HEART DISEASE DATASET USING CLASSIFICATION TECHNIQUES".
- [2] Shan Xu, Haoyue Shi, Xiaohui Duan, Tiangang Zhu, Peihua Wu, Dongyue Liu, "Cardiovascular Risk Prediction Method Based on Test Analysis and Data Mining Ensemble System".
- [3] Theresa Princy. R, J. Thomas, " Human Heart Disease Prediction System using DataMining Techniques", in International Conference on Circuit, Power and Computing Technologies [ICCPCT], 2016.
- [4] Sana Shaikh, Amit Sawant, Shreerang Paradkar, Kedar Patil, " Electronic Recording System-Heart Disease Prediction System", in International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb. 04 – 06, 2015, Mumbai, India.
- [5] Mai Shouman, Tim Turner, Rob Stocker, "USING DATA MINING TECHNIQUES IN HEART DISEASE DIAGNOSIS AND TREATMENT", IEEE, 2012.
- [6] Devendra Ratnaparkhi, Tushar Mahajan, Vishal Jadhav, " Heart Disease Prediction System Using Data Mining Technique ", in International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 08 | Nov-2015, IRJET ISO 9001:2008 Certified Journal Page 1553.
- [7] V. Krishnaiah, G. Narsimha, N. Subhash Chandra, " Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review ", in International Journal of Computer Applications (0975 – 8887) Volume 136 – No.2, February 2016.
- [8] Deepali Chandna, "Diagnosis of Heart Disease Using Data Mining Algorithm", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1678-1680.