

Prediction of Chronic Kidney Disease using Data Mining Classification Techniques and ANN

Sahana B J

Final Year M. tech Pursuing in Computer Engineering,
Department of Computer Sciences & Engineering,
P.E.S College of Engineering Mandya

Abstract - Information mining is a procedure of separating helpful data from gigantic measure of dataset. Information mining has been a present pattern for getting analytic outcomes. In therapeutic application, tremendous measure of unmined information is gathered by the social insurance industry with a specific end goal to find concealed data for successful conclusion and basic leadership. There are numerous information mining systems like grouping, clustering and so on. The goal of our paper is to anticipate Chronic Kidney Disease (CKD) utilizing arrangement methods like Naive Bayes and to foresee the phases of endless kidney illness utilizing the Artificial Neural Network (ANN) like C4.5.

Keywords: Data mining, Classification, Chronic Kidney disease, Naive Bayes, Artificial Neural Network, C4.5.

I. INTRODUCTION

Information mining is a procedure of separating valuable data from enormous measure of dataset. Information mining has been in numerous areas like picture mining, assessment mining, web mining, content mining, diagram mining and so on. The uses of information mining incorporate peculiarity location, monetary information examination, medicinal information investigation, interpersonal organization examination, advertise investigation and so forth. And furthermore it has turned out to be prevalent in wellbeing association as there is a prerequisite of expository strategy for foreseeing and discovering obscure examples. Information mining assumes an imperative part in finding new patterns in social insurance industry. It is especially valuable in wellbeing field when no accessibility of affirmation supporting a specific treatment choice is found. Expansive measure of complex information is being produced by social insurance industry about patients, infections, healing facilities, medicinal equipment's, cases, treatment cost and so forth that requires handling and examination for data extraction. It contains an arrangement of devices and procedures which when connected to prepared information, gives valuable data to medicinal services experts for settling on right choices and that enhances the execution of patient administration errands. Patients with same medical issue can be consolidated together and good treatment arrangements could be given in view of information gathered from past patient.

The worldwide medical issue which has been confronted now days is constant kidney infection (CKD) this is the region of concern. Interminable kidney illness is a condition where kidneys end up plainly harmed and can't channel poisonous materials in our body. Our work dominantly concentrates on identifying life debilitating ailments like Chronic Kidney Disease (CKD) utilizing Classification calculations like Naive Bayes and Artificial Neural Network(ANN) like C4.5 to predicts phases of Chronic kidney disease(CKD).

II. SYSTEM MODEL

Health care industries are providing several benefits like fraud detection in health insurance, availability of medical facilities to patients at inexpensive prices, construction of effective healthcare policies, identification of smarter treatment methodologies, and effective hospital resource management, improved patient care, better customer relation and hospital infection control management. Disease detection is also one of the significant areas of research in medical but still the manual approach is used without automation technique.

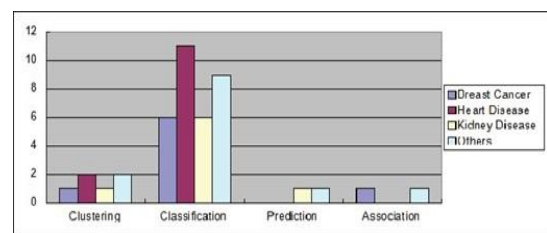


Fig. 1. Data mining techniques used for disease detection

Figure 1 shows a potential use of data mining techniques like clustering, classification which includes DT, Naive Bayes, Neural Network, SVM etc. in predicting heart disease [6], [7], [8], [9], [11], [12], [13], [16], [18], [22], [24]. Classification, association and clustering techniques have also been adopted for breast cancer detection [1], [3], [5], [25], [26], [27], [28], [29]. Other diseases like lung cancer, liver cancer, diabetes, parkinson's disease etc. have also been studied, detected and diagnosed by data mining algorithms [2], [3], [4], [10], [14], [15], [17], [19], [20], [21], [23]. From the fig 1 it shows that the classification technique is more

appropriate than that of other techniques for disease prediction.

The maladies these days are a direct result of the present way of life of individuals, workplace and eating regimen has

Unending Kidney offered ascend to numerous ailments, one of which incorporates constant kidney illness. disease(CKD) is winning these days and has turned into a worldwide medical problem which must be auspicious identified and analysed. Kidneys are essential organs of human body that kill dangerous and undesirable waste from blood bringing about smooth working of body organs. CKD is a condition that depicts loss of kidney capacity after some time making it troublesome for them to channel toxic squanders from the body. Scientists in their current review have tended to the utilization of information digging strategies for CKD location [30], [31], [32], [33], [34], [35].

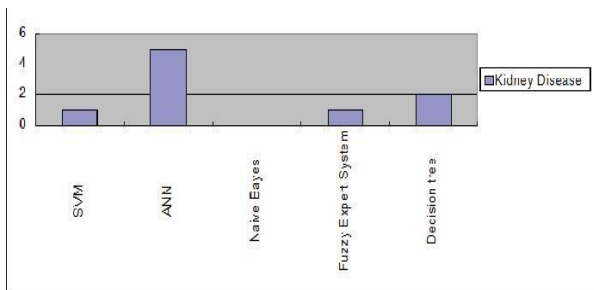


Fig. 2. Classification techniques used for detecting kidney disease stages

It has been observed that classification algorithms have widely been used for identifying and investigating kidney disease. Figure 2 shows that many research work has been conducted using ANN while other techniques like SVM, Fuzzy logic have been used the least. It has also been observed that Naive Bayes has rarely been used. In this research work Naive Bayes algorithm is an important classification algorithm. It is particularly suited when the dimensionality of inputs is high. In this work the dimensionality of dataset is 25.

III. PREVIOUS WORK

Nowadays, health care industries are providing several benefits like fraud detection in health insurance, availability of medical facilities to patients at inexpensive prices, identification of smarter treatment methodologies, and construction of effective healthcare policies, effective hospital resource management, better customer relation, improved patient care and hospital infection control. Disease detection is also one of the significant areas of research in medical. There is no automation for chronic kidney disease prediction.

IV. PROPOSED METHODOLOGY

Chronic kidney infection (CKD) has turned into a worldwide medical problem and is a zone of concern. It is

a condition where kidneys end up plainly harmed and can't channel poisonous squanders in the body. Our work dominantly concentrates on distinguishing life debilitating sicknesses like Chronic Kidney Disease (CKD) utilizing Classification calculations and expectation of chronic kidney ailment stages.

Proposed framework is computerization for endless kidney illness expectation utilizing characterization strategy "Naive bayes" and artificial neural system procedure for stage forecast "C4.5".

In the current work of prediction of chronic kidney disease(CKD) that uses naïve bayes algorithm and C4.5 algorithm . In which the new patient disease prediction depends on the previous patient dataset. This dataset works dynamically.

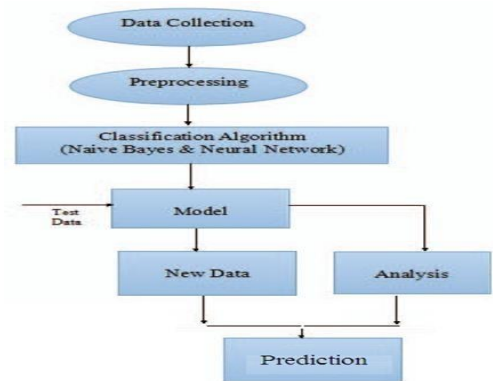


Fig 3 Flowchart of application

Above fig3 shows the experimental setup. The process that involves the collection of patient data from various source then sent to pre-processing to get the data that is chosen then the data is transformed into particular suitable format .Then the data mining technique is applied on the data to extract useful information and the evaluation is done at the end.

V. EXPERIMENTAL RESULTS

The expectation of perpetual kidney sickness and forecast of stages is done utilizing the Naïve bayes and Artificial Neural system C4.5 calculation. This forecast is time effective that procedure time relies on upon the past patient dataset.

Performance Factor

1. Data Structure – array based
2. Memory Utilization – depends on the data set [less for small dataset]
3. No.of.scans – depends on number of diseases
4. Execution time - execution time depends on number of old patien

No of Instances (records)	Execution Time (Secs)
500	3
250	2
100	1
20	0.05

VI. CONCLUSION

This project is a medical sector application which helps the medical practitioners in predicting the disease types based on the symptoms. Patients can also predict diseases by entering symptoms in the form of sentences. It is automation for disease prediction and it identifies the disease, its types and complications from the clinical database in an efficient and an economically faster manner.

It is successfully accomplished by applying the Naïve Bayes algorithm for classification. This classification technique comes under data mining technology. This algorithm takes symptoms as input and predicts the disease based on old patients data.

VII. FUTURE SCOPES

The future work of this research is that we can provide the interaction between doctor, receptionist and admin and also can use sms/email to provide userid and password to the actors.

REFERENCES

- [1] Veenitha kunwar, Khushboo Chandel, A.Sai Sabitha, Abhay Bansal, "Chronic Kidney Disease Analysis Using Data Mining Classification Techniques", [1], 2016.
- [2] Hlaudi Daniel Masethe, Mosima Anna Masethe "Prediction of Heart Disease using Classification Algorithms", [2], 2014.
- [3] N. Aditya Sundar1, P. Pushpa Latha2, M. Rama Chandra, "Performance Analysis Of Classification Data Mining Techniques Over Heart Disease Data Base", [3], 2012.