

A Review of Kidney Prediction using Fuzzy and Neural Network

Gaganpreet Kaur¹ Miss Ada²

²Assistant Professor

^{1,2}Department of Computer Science and Engineering

^{1,2}Sri guru granth sahib world university

Abstract— The huge amount of data generated for prediction of kidney disease are difficult to understand or consisting too many volumes to be processed and analyzed by traditional methods. In the health care industry the data mining is plentiful technique for heap of data. The prophecy of pair of organs in the adomen which excrete urine has been a challenging through and careful search problem for any researcher or technologists. we viewed the fuzzy and neural network techniques for predict kidney diagnose.

Key words: Data mining, kidney disease, Clustering, fuzzy and neural network

I. INTRODUCTION

The main objective of our paper is to learn the different techniques of data mining used in prediction of kidney disease by using data mining tools. Based on data mining techniques many researchers organized their research successfully. Data mining is used for anticipate solution for predicting diagnose such as classification and clustering. Kidney plays a decisive role for destroying the wastage material arrives in the body and provides a equilibrium functionality of the body organs sometimes kidney find out points of dissimilarities as excretion before kidney are unprocurable for clear out the wastage from the body to regulate water and electrolyte poise then the patient have to undergo a dialysis or renal transplantation[1]. Dialysis is artificial process for kidney disease in which kidney injury and cannot purify the renal transplantation is permanent relief for kidney disease[2]. The number of patients on hemodialysis due to end stage kidney disease is increasing. The median survival for these patients is only about 3 years and the cost of providing care is high[3]. The test procedures are assumed to be essential in order to reach the ultimate diagnosis. Else, more number tests could obfuscate the main diagnosis process which may result in trouble in gaining the end results, predominantly in the perceptively of finding disease many tests should be performed [4]. Each kidney contain around a million units called nephrons each of which is Micronics filter for blood. It is possible to loose as much as 90% of kidney function without any experience and symptoms or problem. kidney disease a silent killer[3]. The parameters based on which functionality of kidney is measured are hemoglobin measure, platelet count, urea measure, white blood cell count(WBC), albumin measure, fistula measure, creatinine measure[2]. If this organ doesn't work properly, it can cause many problems for the body and even the human die[5] Approximately 370,000 Americans suffer from end stage kidney disease and require some form of renal replacement therapy (either dialysis or kidney transplantation) to sustain life. Statistics from the United States Renal Data Systems in the year 2000 reported that nearly 250,000 Americans required regular hemodialysis[6]. Symptoms of kidney disease vary depending on the specific type of kidney disease. The symptom of heart disease is change in your urinary, blood in urine, pain or swelling in

back, pain during voiding, nausea and vomiting, or skin rashes.

There are number of factors which increase the risk of kidney disease:-

- Diabetes
- Smoking
- Heart disease
- Male sex
- Drugs abuse/drug overdose
- Ethnicity/Race
- Alcohol intake
- Obesity
- Age
- Family history of kidney disease

II. LITERATURE REVIEW

Govindra.k et al [1]. Describe in their research predict more accurately the dialysis requirement of a patient by using historical and current data of a person by examining kidney functionality. In this paper, they have used rule based prediction method to predict the dialysis requirement of a patient.

B.V Ravindra et al [2]. Choosing appropriate Data mining algorithms and Applying the procedure on Dialysis data set can be taken up as a further study and analysis of survival prediction of a dialysis patient. Also comparative analysis of the results using various data mining tools and assessing is also good one to carry out this work further.

K.R.Lakshmi1 et al[3]. In this research comparison of three techniques which shows ANN algorithm plays a vital role in data mining techniques. ANN shows better results. The performance of ANN shows the high level compare with other two techniques. Hence ANN shows the concrete results with Kidney dialysis of patient records.

DR.S.Vijayarani1 et al[4]. The algorithm which has the higher accuracy with the minimum execution time has chosen as the best algorithm. In this classification, each classifier shows different accuracy rate. SVM has the maximum classification accuracy and it is considered as the best classification algorithm. But Naïve Bayes perform as best with minimum execution time.

Mina Lagzian et al [5]. Mina Lagzian (2012) "A New Online, pp 1097-1101, The final Fuzzy Control Model for Kidney Patients", Scientific Research, Published diagnosis of some diseases depends on many factors and information. Considering all these factors and reviewing all of them is a difficult process, thus providing a mathematical model that can simultaneously consider all these factors is a great help for physicians to diagnose and treat these diseases. In this paper, we propose a new fuzzy control model for kidney transplantation patients. For the inference and conclusion of this model, we use Mamdani approach. Also we employ a new method to smooth the well-known non-

smooth piecewise membership functions, i.e. trapezoidal and half trapezoidal membership functions.

Shital shah et al[6]. The overall classification accuracy for all data mining algorithms was significantly higher using the individual visit dataset over the aggregate data set. The prediction accuracy of individual visit based rule sets increased over the aggregate based rule sets. These improvements were in spite of excluding known significant features such as diagnosis, time for dialysis, target weight, etc. Thus the confidence in the individual data set classifiers is considerably higher. The significant features (identified collectively by data mining) that could be included in the dialysis treatment protocol are diagnosis, time on dialysis, deviation from target weight, blood pressures ranges for different patients, calcium and potassium levels in dialysis solution, total blood volume, blood.

Jamshid Nourozi et al[7] “The effects of the underlying disease and serum albumin on GFR prediction using the Adaptive Neuro Fuzzy Inference System (ANFIS)”, Journal of Health Management & Informatics, Vol:1, Issue:3, pp 46-50, Introduction: Kidney disease is a major public health challenge worldwide. Epidemiologic data suggest a significant relationship between underlying diseases and decrease in Glomerular Filtration Rate (GFR). Clinical studies and laboratory research have shown that the mentioned parameter is effective in development and progression of the renal disease per se. In this study, we used learningbased system based on the neural network concepts.

Rama Devi.E et al[8] “Design Methodology of a Fuzzy Knowledgebase System to predict the risk of Diabetic Nephropathy”, IJCSI, Vol. 7, Issue 5, pp 239-247, The main objective of the design methodology of a Fuzzy. knowledgebase System is to predict the risk of Diabetic Nephropathy in terms of Glomerular Filtration Rate (GFR). In this paper, the controllable risk factors “Hyperglycemia, Insulin, Ketones, Lipids, Obesity, Blood Pressure and Protein/Creatinine ratio” are considered as input parameters and the “stages of renal disorder” is the output parameter. The input triangular membership functions are Low, Normal, High and Very High and the output triangular membership functions are s1, s2, s3, s4 and s5. As the renal complications are now the leading causes of diabetes-related morbidity and mortality, a FKBS is designed to perform the optimum control on high risk controllable risk factors by acquiring and interpreting the medical experts’ knowledge.

Vahid Reza Nafisi et al[9] “Fuzzy Logic Controller for Hemodialysis Machine Based on Human Body Model”, Journal of Medical Signals & Sensors, Jan-Apr; Vol:1, Issue:1, pp 36–48, Fuzzy controllers are being used in various control schemes. The aim of this study is to adjust the hemodialysis machine parameters by utilizing a fuzzy logic controller (FLC) so that patient's hemodynamic condition remains stable during hemodialysis treatment.

Er. Arpit Gupta et al[10] Cluster analysis divides data into groups (clusters) for the purposes of summarization or improved understanding. For example, cluster analysis has been used to group related documents for browsing, to find genes and proteins that have similar functionality, or as a means of data compression. In this chapter we provide a short introduction to cluster analysis. We present a brief view of recent techniques which uses a concept-based clustering approach

A. Data Mining:

Data mining refers to extracting or mining the knowledge from large amount of data. The term data mining is appropriately named as ‘Knowledge mining from data’ or “Knowledge mining”. Data collection and storage technology has made it possible for organizations to accumulate huge amounts of data at lower cost. Exploiting this stored data, in order to extract useful and actionable information, is the overall goal of the generic activity termed as data mining. Data mining is the process of exploration and analysis, by automatic or semiautomatic means, of large quantities of detain order to discover meaningful patterns.

S.NO	Type of disease	Technique	Data mining tool
1	Heart disease	classification	ODND, NCC2
2	cancer	classification	Weka
3	Brain cancer	clustering	k-mean clustering
4	diabetes	classification	ANN
5	HIV/AIDS	Classification, Association rule	Weka 3.6
6	Heart disease	Classification	Weka 3.6.4
7	Heart disease	Classification	tanagra

Table 1: Data Mining Application In Health Care[4][8]:

B. Techniques For Predicting Diagnose:

1) Clustering:

Clustering is a major task in data analysis and data mining applications. It is the assignment of combination a set of objects so that objects in the identical group are more related to each other than to those in other groups. Cluster is an ordered list of data which have the familiar characteristics. Clustering is an important task in data analysis and data mining applications. Clustering is the task of grouping a set of objects so that objects in the same group are more similar to each other than to those in other groups (clusters). Clustering can be done by the different algorithms such as hierarchical- based, partitioning-based, grid-based and density-based algorithms. Hierarchical-based clustering is the connectivity based clustering. Partitioning-based algorithm is the centroid based clustering. Densitybased clusters are defined as area of higher density then the remaining of the data set. Grid based clustering, partition the space into a finite number of cells that form a grid structure on which all of the operations for clustering are performed.[4]. Instead, any desired number of clusters can be obtained by “cutting” the dendrogram at the proper level. Finally, hierarchical techniques are thought to produce better quality clusters. In this section we describe[10]

2) K-Means Algorithm:

K-means algorithm represents each cluster by the mean value of the objects in the cluster .It takes the input parameter, k, and partitions a set of n objects into k clusters so that the resulting intra cluster similarity is high but the inter cluster similarity is low. Cluster similarity is measured in regard to the mean value[2]. Several unsupervised learning algorithms have been proposed which partition the set of objects into a given number of groups according to an optimization criterion. One of the most popular and widely studied

clustering methods is K-means [2]. K-means has a number of variations, depending on the method for selecting the initial centroids, the choice for the measure of similarity, and the way that the centroid is computed. The common practice, at least for Euclidean data, is to use the mean as the centroid and to select the initial centroids randomly.[10]

3) Artificial Neural Network (Ann):

Neural network was traditionally considered as the network of biological neurons. But artificial neural network term are usage in the modern that are in the form of neural network. Artificial neural networks (ANN) have emerged as a result of simulation of biological nervous system, highly sophisticated analytical techniques, capable of modeling extremely complex non-linear functions [3]. ANN mainly consists of three layers that is input layer, hidden layer, output layer. Input layer that is connected to second layer that is called as hidden layer consisting in middle of two layer which is connected to the last one output layer[6]. Raw information is represented as the input layer. Activity of the hidden layer is considered by input layer weight on the connections between input layer and hidden layer. The output layer is depend on the activity of the hidden layer and weight between the hidden and output layer. Back propagation absorbs by iteratively processing a data set of training tuples, comparing the network's prediction for each tuple with the actual. For each training tuple, the weights are modified so as to minimize the mean squared error between the network's prediction and the actual target value.

4) Fuzzy Rule:

Kidney transplantation needs a accuracy in prediction. Most of the researcher proposed different system for kidney diagnose but accuracy in their prediction is not accurate without mathematical logic system[9]. fuzzy logic is useful for chosen prominent decision at the right time for giving various type of treatment which ultimately reduce the condition of being subject to death.[7]. Many researchers works tried to propose mathematical model to control disease adjusting the optimizal dosage of medical to bettwr treatment of disease[8]. Fuzzy reasoning requires some basic rules (inference rules) which can be designed according to expert's decision. Designing these basic rules that are in the form of "if-then", .Accuracy of the results is the prior reason for the development of a new technique for prediction of the kidney diagnose.. The disease prediction needs to be accurate for saving the life of patient.

C. Methodology Used In Data Mining:

Data Mining is core part of Knowledge Discovery Database (KDD). Many people treat Data Mining as a synonym for KDD since it's a key part of KDD process. Knowledge discovery have following steps.

- Data Cleaning - To remove noise or irrelevant data.
- Data Integration - Where multiple data sources may be combined.
- Data Selection - Where data relevant to the analysis task are retrieved from the database.
- Data Transformation - Where data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations.
- Data Mining - An essential process where intelligent methods are applied in order to extract data patterns.

- Pattern Evaluation - To identify the truly interesting patterns representing knowledge based on some interestingness measures.
- Knowledge Presentation - knowledge representation techniques are used to present the mined knowledge to the user.

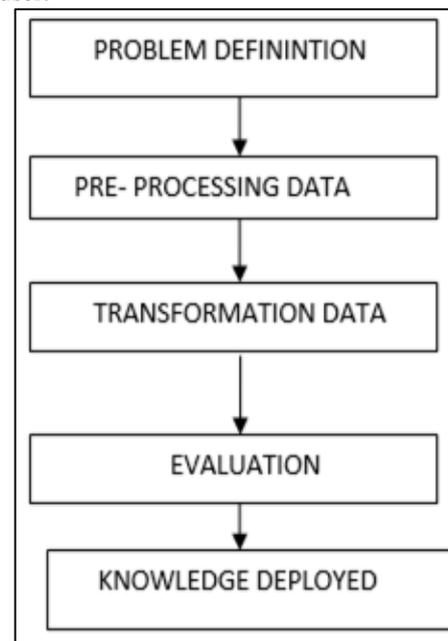


Fig. 1: Methodology of data mining

III. CONCLUSION AND FUTURE SCOPE

The overall objective is to study the various data mining techniques for predicts kidney diagnose. It may be conclude from the above literature review that Artificial Intelligence techniques and fuzzy logic are much more reliable and accurate as compared to conventional statistical methods for kidney disese diagnose. Future scope is in proposing an kidney diagnosis system with important input variables by using combining both fuzzy and neural network technologies called fuzzy and neural network.

REFERENCES

- [1] Govinda.K, Prasanna. S, "Medical Dialysis Prediction using Fuzzy Rules", International Conference on Soft-Computing and Network Security (ICSNS -2015), Feb. 25 – 27, 2015.
- [2] B.V Ravindra, N.Sriram, "Discovery of Significant Parameters in Kidney Dialysis Data Sets by K-Means Algorithm" pg.452-454, 21-22 NOVEMBER 2014.
- [3] K.R. Lakshmi, Y.Nagesh and M.Veera krishna " Performance comparision of three data mining techniques" International journal of asvance in engineering and technology, Vol 7, Issue 1, pp.242–254, Mar,2014.
- [4] DR.S. vijayarani, Mr.s.Dhayanand, "Data mining classification algorithm for kidney prediction", International journal on cybernetic and information, Volume 4, Issue 4, pp.14-24, April, 2015 .
- [5] Mina Lagzian , "A New Fuzzy Control Model for Kidney Patients", Scientific Research, Published Online, vol 3, pg. 1097-1101, March,2012.

- [6] Shital shas, Andrew Kusiak, “Data Mining in Predicting Survival of Kidney Dialysis Patients - Invariant object approach” vol.4049, pp 1-8, JAN 2003.
- [7] Jamshid Nourozi ,“The effects of the underlying disease and serum albumin on GFR prediction using the Adaptive Neuro Fuzzy Inference System (ANFIS)”, Journal of Health Management & Informatics, Vol 1, Issue 3, pg. 46-50, May,2006.
- [8] Rama Devi.E ,“Design Methodology of a Fuzzy Knowledgebase System to predict the risk of Diabetic Nephropathy”, Vol 7, Issue 5, pg 239-247 , September,2010.
- [9] Vaishali Jain (2015) “Improving the Prediction Rate of Diabetes using Fuzzy Expert System”, IJITCS, Vol: 10, pp 84-91.
- [10] Er. Arpit Gupta 1, Er.Ankit Gupta , Er. Amit Mishra “Research paper on cluster techniques of data variations” International Journal of Advance Technology & Engineering Research, Vol.1, issue1, pp 39-47, NOV 2011.

