



DESIGN AND ANALYSIS OF ALGORITHM FOR PREDICTION OF KIDNEY DISEASE

Er. Jyoti Saini
Research Scholar (MTech)
BCET, Gurdaspur

Dr. R.C Gangwar
H.O.D (Associate Professor)
BCET, Gurdaspur

Er. Mohit Marwaha
Assistant Professor
BCET, Gurdaspur

ABSTRACT

Data mining is process of extraction of hidden predictive information from the voluminous databases. It is used to finds previously unknown patterns and trends from databases. These trends are further used for make important decisions, which is based on result of analysis. Day to day basic healthcare industry generates large amount of data. These data have to be analyzed and hidden or meaningful patterns to be discovered. Data Mining techniques are used for disease prediction. This review paper focused on kidney disease by using the classification techniques such as SVM , Neural network and Discriminant analysis.

Keywords : Data Mining, Classification technique, Kidney disease, Dataset, Healthcare, Support vector machine, nural network, Discriminant analysis.

1. INTRODUCTION

Data mining is the analytical process to explore specific data from large volume of data. It is a process that finds previously unknown patterns and trends in databases. This information is further used to build predictive models. The main objective is to learn the different data mining techniques/algorithms which are used in the prediction of Kidney diseases using any data mining tool. Kidney is the most vital part of the human body as life is dependent on efficient working of

Kidney. At any stage of kidney disease, knowledge is power. Knowing the symptoms of kidney disease can help to get the treatment you need to feel your best [1].

The kidneys' functions are to filter the blood. The kidneys are important organs with many important functions in the body, which includes producing hormones, absorbing minerals, and filtering blood as well as producing urine. All the blood in our bodies passes through kidneys several times a day. The kidneys remove wastes, control the body's fluid balance as well as regulate the balance of electrolytes [1]. As the kidneys filter the blood, so they create urine which collects in kidneys' pelvis funnel-shaped structures that drain down the tubes called ureters to the bladder. Each kidney contains around a million units which are called nephrons, each of which is a microscopic filter for blood. It is possible to lose 90% of kidney function without experiencing any symptoms or problems. Kidney disease is a silent killer.

There are number of factors, which increase the risk of Kidney disease:

- Diabetic Patient
- Hypotension
- Smoke
- Tumidity
- Congestive heart failure
- Family history of hydronephrosis
- Large amount of Alcohol consume
- Drug intake
- Age of person
- sex

Symptoms of kidney disease:

- Changes in your urinary function
- Pain or Swelling in the back or sides
- Blood flow in the urine
- Dizziness & Inability to concentrate
- Feeling cold all the time
- Difficulty or pain during voiding
- Extreme fatigue and generalized weakness
- Ammonia breath and metallic taste

- Vomiting
- Skin rashes and itching

Types of Kidney diseases :

- ❖ **Acute renal failure (kidney failure):** Functions of the kidney suddenly fall by this failure. Generally, reason behind this failure is dehydration(blockage in urinary tract) and damage the kidney or may be cause reversible acute renal failure.
- ❖ **Polycystic kidney disease:** Genetic level alteration from a large cysts in both the kidneys and then they damage both the kidneys function.
- ❖ **End stage renal disease (ESRD):** Functions of kidney completely damage by progressive chronic kidney disease and patient with this type of disease require dialysis for survival.
- ❖ **Chronic renal failure:** Functions of kidney partially damage by this failure. Their main causes are diabetes and high blood pressure.
- ❖ **Diabetic nephropathy:** Kidney function damage by high blood sugar from diabetes. Then, ultimately causing chronic kidney disease. Some nutrients such as protein may results in urine.
- ❖ **Hypertensive nephropathy:** Kidney damage caused by high blood pressure. It may eventually result in Chronic renal failure .
- ❖ **Kidney cancer:** Renal cell carcinoma is the most common cancer affecting the kidney and also Smoking is the most common cause of kidney cancer.

DATA MINING

It is main concerned with extracting useful information from large amount of databases. Data mining classification and tools are find unknown patterns or trends from the dataset. The datamining purpose is to automatically find the patterns in the dataset with input and minimal user effort. Data mining's main contribution is in decision making and in forecasting future trends of market. Many organisations use data mining as a tool these days for data analysis as it easily evaluates patterns and trends of market and produce effective results.

DATA MINING APPLICATIONS:

TELECOMMUNICATION INDUSTRY: The telecommunication industry is the most developing industry as it provides several services such as pager, fax, cellular phones and e-mails. Data mining helps to find telecommunication patterns, fraud actions, make superior usage of resources and develop service quality. With the development of computer, telecommunication services have integrated with the communication technologies and works more efficiently.

RETAIL INDUSTRY: In retail industry, data mining is a great application as it collects large amount of data which includes transportation, sales and consumption of goods and services. This data expand rapidly due to increase in purchase and sales in business. Data mining helps to identify customer's buying patterns and trends that lead to improved customer service and customer's satisfaction quality.

FINANCIAL DATA ANALYSIS: Financial data in banking is consistent and of high quality which enables analysis of systematic data in financial industry.. It helps in grouping of customers for target marketing. It also helps in prediction of loan payment and analysis of customer credit policy

HEALTHCARE INDUSTRY: Data mining is very advantageous in healthcare industry in diagnosis of breast cancer, diabetes, kidney and heart diseases. It helps in recognizing patterns and trends in records of patient having similar risk factor and helps in decision making.

2. LITRATURE REVIEW

Over the years, numerous works have been done related to disease prediction system using different data mining algorithms by different authors. They tried to achieve efficient methods and accuracy in finding out diseases related to heart by their work including datasets and different algorithms along with the experimental results and future work that can be done on the system to achieve more efficient results. Different data mining techniques related to disease prediction system has been presented by different authors. In spite of these data, present paper goals at analyzing different techniques of data mining for prediction of kidney disease.

N. SRIRAAM et al [4] presented data mining approach for parametric evaluation to improve the treatment of kidney dialysis patient. Results on the basis of dialysis parameter combination

suggested that its classification accuracy by Association mining is found between the ranges of 50-97.7%. Therefore, such approach can be provide benefits to the clinician and they can easily select the level of dialysis that is basically required for specific patient.

K.R. Lakshmi et al [5] have been reported that performance comparison of Artificial Neural Networks, Logical Regression and Decision Tree are used for Kidney dialysis survivability. Different accuracy measures (classification accuracy, specificity and sensitivity) were selected for the estimation of data mining techniques. They achieved results by using 10 fold cross-validations and confusion matrix for each technique. They found ANN shows better results using Kidney dialysis of patient records.

DSVGK Kaladhar et al [7] described in their research to understand machine learning techniques to predict kidney stones. They predicted good accuracy with C4.5, Classification tree and Random forest (93%) followed by Support Vector Machines (SVM) (91.98%). Logistic and NN has also shown good accuracy results with zero relative absolute error and 100% correctly classified results. For the prediction of accuracy of the data, two curves such as Calibration curves and ROC has been constructed with the help of Naive Bayes. Overall data suggested that machine learning approaches give well defined results for the treatment of kidney stones.

Morteza Khavanin Zadeh et al [8] This research described the prediction of early risk of AVF failure in patients by using supervised classification. Authors used different approaches to predict probability of complication in new haemodialysis patients. But mainly for those patient whom have been suggested by nephrologists to AVF operation.

J. Van Eyck et al [9] After elective cardiac surgery explored data mining techniques for predicting acute kidney injury with Gaussian process & machine learning techniques (classification task & regression task).

3. Proposed Work

Presently various algorithm is available for clustering the pre-processed data, in the existing work they used various classification for disease prediction system and achieved the accuracy. As we can see that there is vast scope of improvement, in our proposed system we will implement SVM Classifier over the Kidney data and will achieve the accuracy more than the present algorithm. To achieve this proposed work user going to follow following major objectives:

- To study the various classification or existing kidney disease prediction algorithm;
- To design a kidney disease prediction algorithm and implement the proposed algorithm.
- To compare the performance of proposed algorithm against the pre-existing kidney disease prediction algorithms.

4. Detail about different classifier's :

Support Vector Machine (SVM) :

For the organization of linear and non linear data a new technique is introduced and need as SVM. The working of this algorithm is given as follows. For the revotation of unique troinig data into higher dimension, It use non linear mapping. Surrounded by this latest dimension, SVM examine for linear optimal repurating hyperplane i.e “decision boundary”. Renting out the tuples of single department to other. Data from one department is always distingished by hypeplane from other department, to achieve largest dimension with suitable mapping. With the help of margins and support vector SVM find hyperplane. In the car of managing complicated nonlinear edge, they perform very perfect and give outstanding performance, Are compare to other technique they are very effective.

Artificial Neural Network (ANN) :

With two methods ANN low the backpropagation ; first by repeatly managing a group of data tuples , rcan d is by performing comparision with their original destination value. This destination value may be previously know as a class label of for classification trouble or forecasting value . Among both the network forcast and destination value, the main motive is to low the weight to lowing the mean square problem. The name back propagation is given, because changes is performed from the output towards the first hidden layer by passing through all hidden layers . This process will continue until weight is finely converges. The major steps that involved are input, output and problem.

5. CONCLUSION

Data mining is very familiar technique to withdrawal the hidden anticipating information from the wide. In this paper, we predict the Kidney diseases using data mining tool. We compare two different Classifiers (SVM and ANN). These two techniques are best in their own ways. It is

found that there is still a lot of future work need to be done. In future these Clusters can more improved by using techniques.

REFERENCES

- [1] Ms. Chaitrali S. Dangare, Dr. Mrs. Sulabha S. Apte, “A data mining approach for prediction of heart disease using neural networks”, international journal of computer engineering and technology”, Vol. 3 , No. 3 , pp. 30-40, (2012).
- [2] M.A. Nishara Banu and B. Gomathy,” *Disease Forecasting System Using Data Mining Methods*”,Vol. 1, No. 4, pp.130-133, (2014).
- [3] Aqueel Ahmed and Shaikh Abdul Hannan,”*Data Mining Techniques to Find Out Kidney Diseases*”, International Journal of Innovative Technology and Exploring Engineering(IJITEE), Vol. 1, No. 4, (2012).
- [4] N.Sriraam, V.Natasha and H.Kaur,”*data mining approaches for kidney dialysistreatment*”, journal of Mechanics in Medicine and Biology, Vol. 06, No. 02,(2006).
- [5] K.R. Lakshmi, Y. Nagesh and M. Veera Krishna,” *Performance comparison of three data mining techniques for predicting kidney disease survivability*”, International Journal of Advances in Engineering & Technology. Vol. 7, No. 1, pp. 242-254,(2014).
- [6] Abeer Y. Al-Hyari,” *Chronic Kidney Disease Prediction System Using Classifying Data Mining Techniques*”, library of university of Jordan, 2012.
- [7] DSVGK Kaladhar, Krishna Apparao Rayavarapu* and Varahalarao Vadlapudi, ”*Statistical and Data Mining Aspects on Kidney Stones: A Systematic Review and Meta-analysis*”, Open Access Scientific Reports, Vol. 1 , No. 12 ,pp. , 2012.
- [8] Morteza Khavanin Zadeh, Mohammad Rezapour, and Mohammad Mehdi Sepehri,” *Data Mining Performance in Identifying the Risk Factors of Early Arteriovenous Fistula Failure in Hemodialysis Patients*”, International journal of hospital research, Vol. 2, No. 1, pp 49-54, (2013).
- [9] J. Van Eyck, J. Ramon, F. Guiza, G. Meyfroidt, M. Bruynooghe, G. Van den Berghe, K.U. Leuven,” *Data mining techniques for predicting acute kidney injury after elective cardiac surgery*”, Springer, (2012).
- [10] Xudong Song, Zhanzhi Qiu, Jianwei Mu,” *Study on Data Mining Technology and its Application for Renal Failure Hemodialysis Medical Field*”, International Journal of Advancements in Computing Technology(IJACT) ,Vol. 4, No. 3, pp. , 2012.
- [11] Abeer Y. Al- Hyari,” *Chronic Kidney Disease Prediction System Using Classifying Data Mining Techniques*”, library of university of Jordan, 2012.
- [12] Luck, M. M., A. Yartseva, G. Bertho, E. Thervet, P. Beaune, N. Pallet, and C. Damon. "Metabolic Profiling of ¹H NMR Spectra in Chronic Kidney Disease with Local Predictive Modeling." In 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), pp. 176-181. IEEE, 2015.

[13] Dr. S. vijayarani , Mr. S. Dhayanand et al . “*Kidney Disease Prediction Using SVM AND ANN Algorithms*”, International Journal of Computing and Business Research (IJCBR), vol. 6 , issue. 2 , pp . 2229-6166 , 2015.

[14] Sunitha Devi. P, Sowjanya. CH, Sunitha. K.V.N, (2014) “*A Review of Supervised Learning Based Classification for Text to Speech System*”, International Journal of Application or Innovation in Engineering & Management, Volume 3, Issue 6, June page no 79-86.

[15] Khyati Chaudhary, Bhawna Mallick, “*Exploration of Data mining techniques in Fraud Detection: Credit Card*”, International Journal of Electronics and Computer Science Engineering, Volume1,Number 3, page no 1765-1771.