International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Survey Paper on a Machine Learning Methodology for Diagnosing Chronic Kidney Disease

Vandana Babu P

UG Student, Department of Master of Computer Science and Engineering, Rajiv Gandhi Institute of Technology , Kottayam, Kerala, India

A B S T R A C T

Kidney is one among the vital organs of the human body, Maintaining overall fluid balance regulation and filtering minerals from blood filtering waste materials from food, and unhealthful substances creating hormones that facilitate turn out red blood cells, promote bone health, and regulate blood pressure. However day by day millions of people are suffering chronic kidney disease .Non-curability and non-detection of this sickness create it more harmful and dangerous as its symptoms are hidden till 90% of urinary organ functions fail. Health industry has taken a move towards mode detection of the CKD at its early stage could be a major and significant task for medical experts. During this scenario computer based mostly intelligent system can be very useful that can diagnose the kidney disease and its stages. The advantage of implementing this technique would be diagnosed at an early stages based on the various symptoms of the patients and thus can help patients to get right treatment and care at right time which will lead to better health. This paper presents a survey on medical expert systems developed for disease diagnosis using different techniques .

Keywords: Machine learning, CKD - Chronic Kidney Disease, GFR - Glomercular Filteration Rate, ESRD - End Stage Renal Disease

1. Introduction

It is explored that kidney disease (KD) enhances the rate of mortality considerably depending on the study by World Health Organization (WHO).Chronic kidney disease is a slowly progressive problem in the kidney which takes time to advance to later stage .Your kidney filter the wastes and excess fluids from your blood which are than excreated in your urine .When CKD reaches an advanced stage
,dangerous level of fluid ,electrolyte and waste can build up in your body .So CKD is an silent and one of the deadly disease in the world. CKD is defined as either the presence of kidney damage or glomercular filteration rate (GFR) <60 mlper minutes for 3 month longer. Normal GFR is about 125ml per minutes and it's reflected by urine clearness measurements and the last stage of the CKD occurs when the GFR is less than 15 ml per minute at this point renal replacement therapy is required. The chronic condition of kidney disease comes when the early stage of the disease is not cured and treated in a serious manner. Although the causes of kidney disease cannot be specified precisely yet there are some issues which causes kidney disease and if they are not cured for a long time they convert kidney disease into chronic kidney disease Diabetes and hypertension are considered as the most common causes of kidney disease. Almost in all cases of the kidney disease, these two are common factors. Around one fourth of all cases of kidney failure are caused by hypertension (high blood pressure) and around one- third of all cases caused by diabetes. Glomerulone phritis and pyelone phritis are also some causes of kidney disease but they are less common as diabetes and hypertension. Some pain killing drugs also cause kidney failure if taken over a long period of time. Fig.1 shows patho-physiology of chronic kidney disease. Kidney disease usually starts silently and discriminating its stages is difficult to some

* *Corresponding author.*
E-mail address: vandanababu007@gmail.com

extent but there are some factors up on which 5 stages are defined although it is not necessary that a kidney patient crosses every stage serially. If GFR is high and no symptoms are seen, kidney's functions are considered normal and kidney is healthy. In stage 1, kidney is damaged with normal kidney function and GFR is above 90. In stage 2, GFR is decreased mildly and it is between 89 and 60. When GFR is decreased 59 to 30, loss of kidney function is moderate; it is considered 3rd stage of Kidney disease. Severe decrease in GFR defines stage 4 when GFR is between 29 and 15 only. Kidney failure is the last stage of chronic kidney disease. It is called 5th stage or ESRD [End Stage Renal Disease] when GFR is below. Kidney transplantation is the best treatment for chronic kidney disease and its success rate is also very good. Approximately 85 to 90 out of 100 cases of kidney transplantation meet success and patients live long with new transplanted organ. It is an operation which replaces diseased or non- functioning organ with a healthy organ donated by living person or a person died and his or her organ is preserved. Kidney transplantation is more effective and better than dialysis treatment. But the expense of kidney transplantation is too heigh it's not affordable for common people.

2. Literature Review

Govinda.K, Prasanna et, al.[1] Medical Dialysis Prediction using Fuzzy Rules, They created an skilled system supported fuzzy rules, with mathematical logic system to infer associate output supported input variable. In a rule based mostly prediction technique a group of rules area unit accustomed predict the result supported a given set of knowledge. Rule based method is the most appropriate method to determine the result as the prediction is concerned with the life of a patient; it can't be done on basis of probabilistic theory. Here, they are defined a set of if-then rules along with and-or mechanism for the purpose of rule based prediction and the membership function used is triangular. But the problem of this method is it is not accurate, so the result are perceived based on the assumption, so it may not be widely accepted . Fuzzy system doesn't have the potential of machine learning further as neural network

Dr. Uma N Dulhare Professor, Mohammad Ayesha et.al [2] Extraction of Action Rules for Chronic KidneyDiseaseusingNaiveBayesClassifier,theyplannedthework,ExtractionofActionRulesforChronic Kidney Disease using Naive Baye's Classifier ,they are not only extracting action rules based on stages but also predicting CKD by using naive baye's with One-R attribute selector which helps to prevent the advancing of chronic renal disease to further stages. In Naive baye's with one-R is also known as Bad estimator. Naive Baye's Classifier are a family of straight forward "probabilistic classifier "based on applying baye's theorem with robust independence assumptions between options. The Bayesian classifier represents a supervised learning methodology and for classification bayesian statistical procedure is employed. An underlying probabilistic model and it permits us to capture uncertainty regarding the model in a principled way by determining probabilities of the outcomes. One of the disadvantages of this approach is that since it is an probabilistic model it is not suitable for analyzing the attributes of the dialysis dataset it predicts using probability. This model gives moderate results in prediction. The main limitation of Naïve Bayesis Zero probability problem, that the event will not happen, every point has zero probability but every point can be possible outcomes.

Jiongming Qin,Lin Chen[3]Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms Chongqing ,They proposed this work to determine the kidney function failure by applying the classification algorithm on the test result obtained from the patient medical report and to reduce the diagnosis time and to improve the diagnosis accuracy using classification algorithms namely Back- Propagation Neural Network, Radial Basis Function and Random Forest. Accuracy calculated by RBF(85.3 %) is highest followed by Back-Propagation (80.4%) and Random Forest (78.6%). They concluded that from the experimental result, the Radial Basis Function is the better accuracy for predicting chronic kidney disease

Adeola Ogunleye and Qing-Guo Wang et al.[4] present a XGBoost Model for Chronic Diagnosis ,they used several typical and recent AI algorithms in the context of CKD and the extreme gradient boosting (XGBoost) is chosen as a base model for its high performance. The reduced model using fewer options is desirable while it should still maintain high performance. The set-theory primarily based rule is presented which combines a few feature selection methods with their collective strengths. The XG Boost method has been studied and optimized for CKD identification. Three feature selecting techniques are combined by leveraging the strengths of each technique. These models coupled with the experience of a nephrologist can help reduce the cost and time to diagnose a CKD patient. It is noted that the some tests for CKD may manufacture pictures as raw data. The image based feature extraction and learning show great potentials in ML in general and medical application in particular. XG-Boosting is that it is sensitive to outliers since each classifier is obligated to repair the errors in the predecessors. Thus, the strategy is just too dependent on outliers. Another drawback of this method is almost impossible to scale up. This is because every estimator bases its correctness on the previous predictors, thus making the procedure difficult to.

Hodneland et al.[5] proposed a Vivo Detection of Chronic Kidney Disease Using Tissue Deformation Fields From Dynamic MR

Imaging , He explore the utilization of image registration methods for detecting renal morphologic changes in patients with CKD. They propose a framework for detecting pathological changes in stiffness using image registration. And there for the findings are based in a small cohort, and want to be verified during a larger clinical study. The simulation study indicated that image registration has a borderline sensitivity close to the stiffness changes seen in delicate CKD. Still, the algorithms were capable of showing significant correlations with biopsy assessments, and future implementations will demonstrate whether image registration has sufficient abilities to be used in personalized medicine as a complementary tool for quantification of gentle pathological changes in tissue stiffness. They found that the absolute deformation, normalized volume changes, yet as pressure gradients correlated significantly with arteriosclerosis from biopsy assessments. But current image registration methodologies are lacking sensitivity to recover mild changes in tissue stiffness. Thus this model may be more developed in terms of sensitivity and specificity, it will offer clinicians with a non-invasive tool of high spatial coverage available for characterization of arteriosclerosis and potentially other pathological changes observed in chronic kidney disease. But it's very costly depending on the system used, the number of detectors purchased and also very time consuming

Srinivasa R. Raghavan, et.al [6] Developed a decision Support for qualitative analysis Treatment of Chronic excretory organ Failure, they describe the development of associate degree intelligent system building on existing information and prone to reconfiguration on the premise of data non heritable throughout the employment of the system. Varied call support techniques were wont to style and develop the choice support modules. During this paper

they in brief reviews the literature on clinical support, and sets for support techniques employed in developing a qualitative analysis decision network. The DARWIN call support style method demonstrates the feasible mess of implementing versatile clinical call support exploitation commonplace software package tools. The DARWIN call network was developed exploitation Microsoft Active Server Pages, Visual Basic, and Microsoft SQL Server because the knowledge base. They used image application development methodology for developing the choice network.

Jianliang Gao- Ling Tian et .al [7] proposed a work of study the similarity of diseases and how it help us to explore the pathological characteristics of composite diseases, and help provide dependable reference information for inferring the relationship between unknown diseases and known diseases, so as to make an effective treatment plans. To obtain the similarity of the disease, most previous methods either use a one similarity metric like semantic score, functional score from single data source, or utilize weight coefficients to easily mix multiple metrics with completely different dimensions. During this paper, they proposes a idea to predict the similarity of diseases by node representation learning. They first integrate the semantics core and topological score between diseases by combining multiple knowledge sources. Then for every disease, its integrated scores with all alternative diseases are utilized to map it into a vector of the same spatial dimension, and therefore the vectors are used to measure and overall analyze the connection between diseases. Lastly, perform a comparative experiment supported on benchmarks and different other disease nodes outside the bench mark set. Using the statistics like average, variance, and coefficient of variation in the bench marks to evaluate multiple methods demonstrates the effectiveness of this approach with in the prediction of comparable diseases. This has three main disadvantages, First, the approaches are inherently transductive and don't generalize to unseen nodes and alternative graphs. Second, they're not space-efficient as a feature vector is learned for every node that is impractical for giant graphs. Third, most of those approaches lack support for attributed graphs.

V Ganapathi Raju1, K PrasannaLakshmi, K. Gayathri Praharshitha, Chittampalli Likhitha et.al[8], Prediction of chronic kidney disease (CKD) using Data Science. The main motive of this work is to establish the existence of chronic kidney disease by imposing various classification algorithms on the patient medical record. This analysis work is primarily concentrated on finding the simplest appropriate classification formula which might be used for the diagnosis of CKD based on the classification report and performance factors. Empirical work is performed on completely different algorithms like Support Vector Machine, Random Forest, XG-Boost, Logistic Regression, Neural networks, Naive Bayes Classifier. The experimental results show that Random Forest and XGBoost offer better results compared to other classification algorithms. The objective of this analysis is to observe classification algorithmes to analyses and predict CKD. We have compared the performance of five classifiers within the prognosis of CKD. The experimental results of our proposed technique have demonstrated that RF and XGB have produced superior prediction performance in terms of classification accuracy for our considered dataset. But when we using the collected data, The data could be incomplete, Missing values, even the shortage of a section or a considerable part of the information, may limit its usability. The information collected from different sources can vary in quality and format. Data collected from such diverse sources as surveys, e-mails, data-entry forms, and the company website can have completely different attributes and structures. Data from various sources might not have abundant compatibility among data fields

RC Jalandhar, Dr. Sheetal Kalra [9] Disease Prediction using Hybrid K-means and Support Vector Machine , Handling this vast amount of data in medical field is challenging, so there is a need to mine this data in order to extract useful patterns for disease prediction. A hybrid K-means and Support Vector Machine algorithm for disease prediction is proposed in this paper. The proposed hybrid K- means algorithm is helpful in choosing initial centroids, number of clusters and also to improve the efficiency of K-means algorithm. The hybrid K-means algorithm is used for dimensionality reduction of the dataset which is given as an input to Support Vector Machine classifier. The simulation is performed in MATLAB and from the results it has been analysed that the accuracy of the classification is improved and the processing time to obtain the final output is reduced. A hybrid K-means and Support Vector Machine algorithm for disease prediction is proposed in this paper to improve the efficiency and accuracy for prediction. The initial centroids are randomly selected in case of simple K-means algorithm but it is not so in proposed algorithm. The proposed work is to select the initial centroids by partitioning the data into k equal parts. The simulation is performed on diabetes dataset in MATLAB. The ultimat results of simulation shows that the potency achieved by planned algorithmic rule is higher than easy K-mean algorithmic rule. The ultimate results of simulation shows that the accuracy achieved by purposed algorithmic rule is best than easy K-means algorithmic rule. In SVM selecting a "good" kernel operate isn't straightforward and long coaching time for giant datasets.Troublesome to grasp and interpret the ultimate model, variable weights and individual impact.

J.R. Cuevas, E. L. Dom´ınguez and Y . H. Vel´azquez et .al [10] Telemonitoring System for Patients with Chronic Kidney Disease Undergoing Peritoneal Dialysis, In this work, they proposed the analysis, design and implementation of a telemonitoring system, to carry out a continuous control and remote monitoring of patients with CKD undergo a Peritoneal Dialysis (PD) treatment. The proposed system consists of a native app for Android, where the patient records his dialysis baths and other relevant data, and a mobile web app in which the medical staff can monitor the patient condition and give recommendations. Some core services our system provides are: data record either about CAPD (Continuous Ambulatory Peritoneal Dialysis) or APD (Automated Peritoneal Dialysis)dialysis sessions, alerts generation by detecting biomedical indices outside ranges established by the doctor, sending emails and text messages with the patient's location and details of why that warning occurred, report generation and data query. Is not accessible for everyone. And requires good broad band connectivity, which is hard to achieve for small healthcare institutions and rural hospitals.

Jayson Mc Allister, Zukui Li, Jinfeng Liu , and Ulrich Simons at.al [11] Erythropoietin Dose Optimization for Anemia in Chronic Kidney Disease Using Recursive Zone Model Predictive Control, In this paper, a novel approach to erythropoietin dosing for anemia management in patients with chronic kidney disease is introduced. A weighted constrained quadratic programming problem is proposed to recursively estimate patient models providing better estimation results than standard system identification methods. The constrained optimization modeling method provides a stable autoregressive with exogenous inputs model that mimics parameter patterns learned from a more complex nonlinear physiologically based model. Zone model predictive control (ZMPC) is compared with standard MPC, digital PID controllers, and economic ZMPC for computer simulations using a designed patient simulator. The results show that weighted recursive ZMPC offers the most desirable controller characteristics while offering significant improvements over the proprietary anemia management protocol used for comparison in the computer simulations. This paper proposes a novel modeling and control algorithm, based on recursive weighted constrained ARX modeling and ZMPC, for the regulation of hemoglobin concentration in patients with CKD undergoing ESA treatment. A patient simulator is designed based on an existing nonlinear PK/PD model and is used to test the modeling and control

algorithms in computer simulations. Unlike the economic model predictive controller, the set point tracking model predictive controller, and the digital PID controllers tested, the authors show that the proposed method and a single set of tuning parameters can be used that still provides robust control performance on a large number of different simulated patients. In the computer simulations, the proposed modeling and control algorithm significantly reduces medication consumption while also improving the number of hemoglobin measurements within the target zone as compared with a population-based AMP typically used within a clinical setting. But it is less tolerant since a merit function is still used. And also installation and maintenance is very expensive.

3. Conclusion

Chronic kidney disease is one of the deadly diseases spreading all over the world and manually diagnosis of this disease is very difficult since most of its symptoms are common to other diseases. A number of medical and computer science scholars had developed computer based intelligent system or medical diagnosis system for chronic kidney disease diagnosis. In this paper a survey is made for chronic kidney disease diagnosis expert system developed by the persons, the methodology they adapted, the dataset they used, accuracy of the system and their contribution to the real world. This survey explores that a lot of work is yet to be done in the field of developing expert system for Chronic kidney disease diagnosis using authorized dataset like UCI dataset machine learning repository and a powerful logic like fuzzy logic. In the future work , to implement the work left by the previous researchers.

REFERENCES

[1] Govinda.K,Prasanna.“MedicalDialysisPredictionusingFuzzyRules”InternationalConferenceon Soft-Computing and Network Security (ICSNS -2015), Feb. 25 –27, 2015’.

[2] Dr. Uma N Dulhare,Mohammad Ayesha, “Extraction of Action Rules for Chronic Kidney Disease using Naive Bayes Classifier” IEEE International Conference on Computational Intelligence and Computing Research2016.

[3] JiongmingQin,LinChen,YuhuaLiu,ChuanjunLiu,ChanghaoFeng”MachineLearningMethodology for Diagnosing Chronic Kidney Disease, College of Electronic and Information Engineering, Kyushu University, Fukuoka 819-0395, Japan,2019.

[4] Adeola Ogunleye and Qing-Guo Wang”XGBoost Model for Chronic Kidney Disease Diagnosis” IEEE/Acm Transactions On Computational Biology And Bioinformatics, Vol. 17, No. 6, November/December 2020.

[5] E. Hodneland, E. Keilegavlen, E. A. Hanson, E. Andersen, J. A. Monssen, J. Rorvik, S. Leh, H.-P. Marti, A.Lundervold, E.Svarstad, andJ.M.Nordbotten,“InVivodetectionofchronickidneydisease using tissue deformation fields from dynamic MR imaging,” IEEE Trans. Biomed. Eng., vol. 66, no. 6, pp. 1779–1790, Jun. 2019.

[6] Srinivasa R. Raghavan, Member, IEEE, Vladimir Ladik, and Klemens B. Meyer” Developing Decision Support for Dialysis Treatment of Chronic Kidney Failure ”IEEE Transactions On Information Technology In Biomedicine, VOL. 9, NO. 2, JUNE 2005.

[7] Jianliang Gao , Ling Tian , Jianxin Wang, Yibo Chen, Bo Song , and Xiaohua Hu ”Similar Disease Prediction With Heterogeneous Disease Information
Networks ” IEEE Transactions On Nano Bioscience, VOL. 19, NO. 3, JULY 2020.

[8] N V Ganapathi Raju1, K Prasanna Lakshmi2, K. Gayathri Praharshitha, Chittampalli Likhith, ”Predic-tion of chronic kidney disease (CKD) using Data Science.”Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2019.

[9] Sandeep Kaur,Dr. Sheetal Kalr ”Disease Prediction using Hybrid K-means and Support Vector Machine” Department of Computer Science Guru Nanak Dev University, RC Jalandhar, Punjab,2017..

[10] J.R.Cuevas,E.L.Dom´ınguezandY.H.Vel´azque”TelemonitoringSystemforPatientswithChronic Kidney Disease Undergoing Peritoneal Dialysis ”IEEE Latin America Transactions, Vol. 14, No. 4, April 2016.

[11] Jayson McAllister, Zukui Li, Jinfeng Liu , and Ulrich Simons ”Erythropoietin Dose Optimiza-tion for Anemia in Chronic Kidney Disease Using Recursive Zone Model Predictive Control”, IEEE Transactions On Control Systems Technology, Vol. 27, No. 3, May 2019Van der Geer, J., Hanraads, J. A. J., & Lupton.