ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 7, July 2021

A survey on Support Vector Machine in Chronic Kidney Disease Prediction

[1] Pallavi Sharma, [2]Dr. Gurmanik Kaur
[1] Ph.D. Scholar, Department of ECE, [2] Associate Professor, Department of EE, SBBSU, Padhiana, Punjab, India

Abstract: Chronic kidney disease (CKD) is one of the main reasons behind death all throughout the world these days. The term "chronic kidney disease" signifies enduring damage to the kidneys that can deteriorate over the long run. On the off chance that the damage is very terrible, then kidney may quit working. This is called End stage renal failure. The prediction of CKD is perhaps the most significant and challenging issues in medical services examination. To acquire the hidden data from the given dataset, data mining is utilized to settle on the decisions. Big data is the latest advancement used to store and deal with the voluminous data and that data can be organized data, unstructured data and semi-organized data. This paper aims to assist in the prediction of chronic kidney disease (CKD) by utilizing the support vector machine (SVM) classifier in medical domain.

Keywords: Chronic Kidney Disease, Data mining, Support vector machine

I. INTRODUCTION

Amount of data in the medical field is expands day by day. It is truly challengeable undertaking to handle this large quantity of data and concentrate them to acquire useful information for successful choice making. That is the reason medical industry thinks of new arrangement as applying data mining methods which will give productive choice from tremendous database. This data mining method is valuable to take care of different real world issues. However, medical care industry is one of the field giving biggest measure of data collected utilizing diverse clinical reports and patient history's [1].

Data mining manages the extraction of helpful information from voluminous data. Presently, numerous medical databases have made on the premise of the advancement of the medical-care database management system. Complex medical data can effectively be kept up by using data mining methods. Data mining has found or either emerged the non-identical direction in the medical-care services association. Data can be accessible in numerous formats like text, signals or graphics [2].

Different data mining classification techniques and machine learning algorithms are applied for prediction of chronic diseases. Here Chronic kidney disease (CKD), too known as chronic renal disease, is an abnormal function of kidney or a reformist disappointment of renal function over a time of months or a long time. Frequently, chronic kidney disease is analyzed subsequently of screening of individuals known to be in danger of kidney issues, such as those with hypertension or diabetes and those with a blood relative with CKD. It is separated from intense kidney disease in that the decrease in kidney function should be available for more than three months. This work predominantly centered around, prediction of chronic kidney disease. CKD is predicted using classification techniques of data mining [3]. This paper prevalently centers on foreseeing life compromising disease like CKD utilizing classification algorithm as Support Vector Machine (SVM).

SVM is machine learning, directed learning algorithm on the foundation of factual learning ideas. SVM has the superior capacity to anticipate, investigations, relapse and groups dataset [4]. It produces a different hyper plane in the descriptor interplanetary of the training data and blends are grouped dependent on the transversely of hyper plane found. It is utilized to predict and examine the dataset relapse and classification strategies [5]. SVM is directed learning algorithms which are for the most part used data mining classification. SVM provides the right outcome by partner other classification techniques. By amplifying the joined between the cases of two classes, it can limit the error. The advantage of the SVM is that by utilization of kernel trick, the distance between a molecule and the hyper plane can be determined in a changed (nonlinear) feature space, lacking the explicit change of the first descriptors.

Literature Review

Researchers have led various examines identified with CKD utilizing diverse data mining techniques lately.

Chronic kidney disease has become a worldwide medical issue and it's an area of concern. It's a situation where kidney's become harmed and cannot filter harmful materials in the body.

Vijayarani and Dhayanand [6] presented a classification process that is used to classify four types of kidney diseases. Correlations of SVM and Naive Bayes (NB) order calculations are set up based on performance factors, classification, accuracy and execution time. Thus, the SVM accomplishes better classification execution. Subsequently it is considered as the best classifier when contrasted with NB classifier algorithm.

Shetty et al. [7] utilized SVM, K-nearest neighbor (KNN) data mining to predict CKD using clinical information. Higher upsides of the chosen performance indices (accuracy=90.09 %, recall=1, precision=0.5) suggested that the SVM classifier performed better compared to the KNN.

Tekle et al. [8] analyzed 14 different attributes associated to CKD patients and estimated accuracy for the Decision Tree (DT) and SVM algorithm. From the results analysis, it was noted that the accuracy obtained by the DT algorithm is 91.75 % and the SVM is 96.75 %.

Kaur and Sharma [9] compared the performance of KNN and SVM classifier for predicting the CKD. The results revealed that the SVM classifier had a higher accuracy (78.09%) and lower error rate (21.9%) when compared to the KNN classifier.

Vijayarani and Dhayanand [10] used SVM and Artificial Neural Network (ANN) to classify kidney diseases such as Acute Nephritic Syndrome (ANS), CKD, Acute Renal Failure (ARF), and Chronic Glomerulonephritis (CG) using synthetic kidney function test (KFT) dataset. The classification accuracy results (76.32 %) revealed that SVM outperformed ANN.

Charleonnan et al. [11] utilized the DT, logistic regression (LR), SVM and KNN as CKD detection classifiers by utilizing UCI machine learning repository CKD dataset. The results showed that the SVM technique performed better as far as detection accuracy and sensitivity. Four classifiers average accuracy was tried five times. As indicated by the experimental results, the SVM classifier created the highest accuracy of 98.3%, while the LR, DT, and KNN delivered average accuracy of 96.55%, 94.8%, and 98.1%, respectively.

Chen et al. [12] presented three multivariate models, KNN, SVM, and soft independent modeling of class analogy (SIMCA), to evaluate patient risk utilizing clinical data from the UCI machine learning repository. Various forms of composite data were likewise used to evaluate the feasibility and robustness of these models in CKD hazard assessment, in which relative unsettling disturbances were acquainted with mimic measurement variances brought about by environmental and instrument noises. For the original data set, the three proposed multivariate models recognized between CKD and non-CKD patients with generally accuracies of over 93%. In this study, KNN and SVM perform better compared to SIMCA.

conclusion

The objective of this study is to observe SVM classifier to analyze and predict CKD. It is seen that results may vary for chronic kidney disease diagnosis based on the different data mining tools and techniques used. Data mining gives good results in disease diagnosis when appropriate tools and techniques applied. Many researchers have been led different data mining, classification algorithm like KNN, ANN, NB, SVM, Decision tree (J48, C4.5) and feature selection to improve the accuracy of the algorithm. But SVM is considered as the best classifier when contrasted with other classifier algorithms.

REFERENCES

[1] Neelamadhab Padhy, Dr. Pragnyaban Mishra and Rasmita Panigrahi, "The Survey of Data Mining Applications and Feature Scope," International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), 2(3), pp 43-58, 2012.

[2] J.C. Prather, D. F. Lobach, L. K. Goodwin, J. W. Hales, M. L. Hage, and W. E. Hammond, "Medical data mining: knowledge discovery in a clinical data warehouse," Proceedings of the AMIA Annual Fall Symposium, pp 101–105, 1997.

[3] Subhash Chandra Pandey, "Data mining techniques for medical data: A review," International Conference on Signal Processing, Communication, Power and Embedded System, Paralakhemundi, India, pp 972-982, 2016.

[4] H. Polat, H. D. Mehr, and A. Cetin, "Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods," J. Med. Syst., Vol. 41, no. 4, Apr. 2017.

[5] K. Chandel, V. Kunwar, S. Sabitha, T. Choudhury, and S. Mukherjee, "A Comparative study on thyroid

disease detection using K-nearest neighbor and Naive Bayes classification techniques," CSI Trans. ICT, Vol. 4, no. 2–4, pp. 313–319, Dec. 2016.

Vijayarani, S. and Dhayanand, S., "Data mining classification algorithms for kidney disease prediction", Int. J. Cybern. Inf. (IJCI), Vol. 4, No. 4, pp 13–25, 2015.

A. R. Shetty, F. B. Ahmed, V. M. Naik, "CKD Prediction Using Data Mining Technique as SVM and KNN with Pycharm", International Research Journal of Engineering and Technology, Vol. 6, No. 5, pp 4399-4405, 2019.

S. Tekale, S. Pranjal, S. Wandhekar, A. Chatorikar, "Prediction of Chronic Kidney Disease Using Machine Learning Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 7, No. 10, pp 92-96, 2018.

G. Kaur, A. Sharma, "Predict Chronic Kidney Disease using Data Mining Algorithms in Hadoop", International Journal of Engineering Researches and Management Studies, Vol. 5, No. 2, pp 34-48, 2018.

S. Vijayarani, S. Dhayanand, "Kidney Disease Prediction using SVM and ANN Algorithms", International Journal of Computing and Business Research, Vol. 6, No. 2, 2015.

A. Charleonnan, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwannawach, N. Ninchawee, "Predictive Analytics for Chronic Kidney Disease using Machine Learning Techniques", IEEE International Conference on Management and Innovation Technology, 2016.

Z. Chen, X. Zhang, Z. Zhang, "Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate models", Int. Urol. Nephrol., Vol. 48, No. 12, pp 2069–2075, 2016.