IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Special Issue 1, February 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Special Issue 1, February 2023 ||

Early Detection of Chronic Kidney Disease Using Machine Learning

Mr.P.Manivannan[1], M.Prabhu[2], B.Kannan[3] , S.Harishkumar[4], M.Lokaprabhu[5]

Associate Professor, Department of Electronics and Communication Engineering, Adhiyamaan College of Engineering, Hosur, Krishnagiri, Tamil Nādu, India. [1]

U.G. Student, Department of Electronics and Communication Engineering, Adhiyamaan College of Engineering, Hosur, Krishnagiri, Tamil Nādu, India. [2-5]

ABSTRACT: **-** The field of life sciences has developed massively and generated a large amount of information from electronic medical records. As a result, there is an urgent need to generate knowledge from this vast amount of data. Data mining techniques and machine learning play a major role in this area of life sciences. Chronic kidney disease (CKD) is a condition in which the kidneys are damaged and unable to filter blood as they normally would. A family history of kidney disease or kidney failure, high blood pressure, and type 2 diabetes can lead to CKD. I have. This is permanent damage to the kidneys and is likely to worsen over time. Very common complications resulting from kidney failure include heart disease, anaemia, bone disease, high levels of potassium and Calcium. At worst, this can lead to complete renal failure and a life-threatening kidney transplant. Early detection of CKD can greatly improve quality of life. This requires good predictive algorithms to predict CKD at an early stage. The literature presents a wide range of machine learning algorithms used to predict CKD. In this paper, we use data pre-processing, data transformation and various classifiers to predict his CNI, and also propose an optimal prediction framework for CNI. Our framework results show promising results for better prediction in the early stages of CKD.

KEYWORDS: chronic kidney disease, decision tree, machine learning, random forest, supportvectors**.**

I. INTRODUCTION

Chronic kidney disease ranks 18th and affects 10% of the world's population, according to the list of total deaths from fatal illnesses published by the Global Burden of Disease survey in 2010 [1]. Approximately 37 million adults have CKD to date, with millions of others at high risk [2]. CKD has no age limit and can occur at any age. Additionally, if someone already has CKD, the likelihood of a rapid deterioration of kidney function increases.

Procrastination can lead to serious kidney damage. Therefore, early detection of this disease is very important for successful treatment. However, CKD has no subjective symptoms in the early stages, so there is no choice but to perform tests. 753 million people have CKD. Worldwide in 2016 he had, deaths. This includes more than 1 million of his people who died of kidney failure in over 100 developing countries because they couldn't afford renal replacement therapy or a kidney transplant. The earlier the diagnosis is made, the more likely it is that the disease will be slowed down. Rapid detection increases the likelihood of slowing or halting progression in the early stages. Applying machine learning prediction algorithms is a great idea to detect the occurrence of CKD. In this research study, he took some widely used machine learning algorithms from several articles and journals and applied them to predict CKD. The proposed study tries to find the best algorithm for predicting CKD by comparison. There are several algorithms such as RF, DT, ADB, GB, KNN, XGB, GNB, EXT and their performance was meticulously determined using performance metrics such as accuracy, accuracy and recall f1 score . We also used voting classifiers, cross-validation, AUC scores, and ROC scores to measure the efficiency of different classifiers. This paper is organized as follows. We first mentioned the dataset and explained the motivation for the experiment. Next, we describe the proposed method and classification algorithm, and describe the details and results of the experiment. Finally, I've outlined some plausible ways to further improve the work.

II. RELATEDWORK

Data mining techniques and machine learning play an important role in this area of life sciences. Chronic kidney disease (CKD) is a condition in which the kidneys are damaged and unable to filter blood as they normally would. A family history of kidney disease or kidney failure, high blood pressure, and type 2 diabetes can lead to CKD. I have. Early detection of CKD can greatly improve quality of life. This requires better predictive algorithms that predict CKD at an earlier stage. The literature presents a wide range of machine learning algorithms used to predict CKD. In this paper, we predict his CKD using data preprocessing, data transformation, and various classifiers, and also propose an optimal prediction framework for CKD. Framework results show promising results for better prediction in the early stages of CKD.2.H. A. Wibawa.et al[2] proposed and evaluated a kernel-based extreme learning machine (ELM) for predicting chronic kidney disease. We compare the performance of four kernel-based ELMs, namely RBF-ELM, Linear-ELM, Polynomial-ELM and Wavelet-ELM, with that of a standard ELM. we demonstrate how eight different machine learning (ML) algorithms can be used to instantly detect CKD infections given patient health record information. We use a hospital-provided dataset of approximately 2 months in this period to identify the likelihood of chronic kidney disease. This study used Extra Tree Classifier (EXT), AdaBoost (ADB), K Nearest Neighbors (KNN), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), Decision Tree (DT), Gaussian Naive Bayes (GNB) Did. and Random Forest (RF) to get the best prediction results. After preprocessing the data, this study applied ML algorithms to compare their performance and finally got accurate results. Performance is analyzed using F1 Score, Precision, Accuracy, Recall and AUC Score. Our analysis results show that K-Nearest Neighbors and Extra Tree Classifier outperform other algorithms, achieving 98% accuracy of 99% before Gradient Boost.

.

III. METHODOLOGY

Using certain characteristics such as total bilirubin, direct bilirubin, alkaline phosphatase, total protein, albumin, and globulin, this software can determine whether a patient has liver disease or not. It is necessary to use supervised le to resolve this binary classification issue. Each data point has ten attributes, and there is a label that indicates if the patient has liver disease or not.In order to find the answer, our goal should be to train a variety of supervised learning models on this dataset in order to create a high-performing model that can accurately identify any new data point as positive or negative and outperform the benchmarks. In this project, we gather data from a data set, and the health specialist can enter the data for testing using our web application. In this application, we perform data cleaning and pre-processing, extensive data analysis, data visualization, and machine learning using supervised learning algorithms, decision trees, K nearest neighbor's, logistic regression, and support vector machines. This approach predicts a person's liver status based on variables such as total bilirubin, direct bilirubin, albumin, and total protein. In addition, donor and recipient information required for liver and blood transplants can be added to the system.

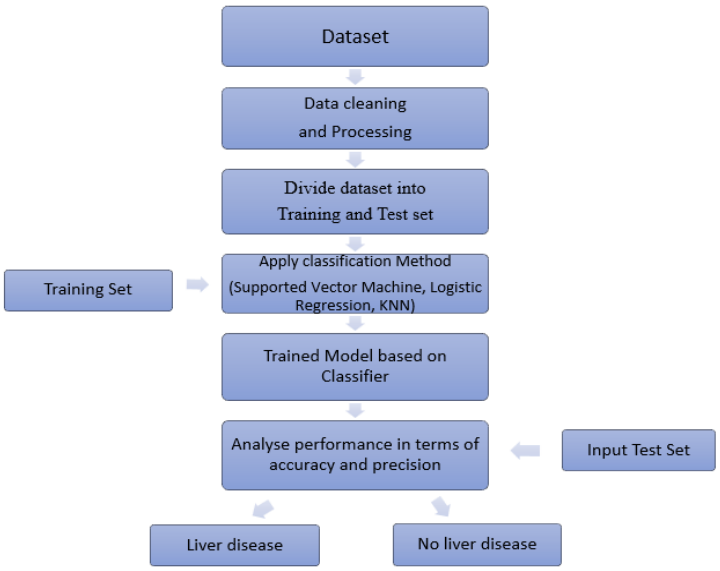

Fig 1:Block Diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Special Issue 1, February 2023 ||

Support vector machines, or SVMs, are one of the most popular supervised learning algorithms used for both classification and regression problems. However, it is mainly used for machine learning classification problems.
The goal of the SVM algorithm is to create optimal lines or decision boundaries that can divide the n-dimensional space into classes so that new data points can be easily placed in the correct category in the future. This optimal decision boundary is called a hyperplane. SVM chooses extrema/vectors to help create hyperplanes. These extreme cases are called support vectors, and the algorithm is called a support vector machine. Support Vector Machines (SVM) is supervised machine learning. Algorithm used for both classification and regression. Also known as a regression problem, it is best suited for classification. The goal of the SVM algorithm is to find a hyperplane in the N-dimensional space that uniquely classifies the data points. The dimension of the hyperplane depends on the number of features. If the number of input features is 2, the hyperplane is just a line. If the number of input features is 3, the hyperplane will be a 2D plane. It becomes difficult to imagine when the number of features exceeds three.

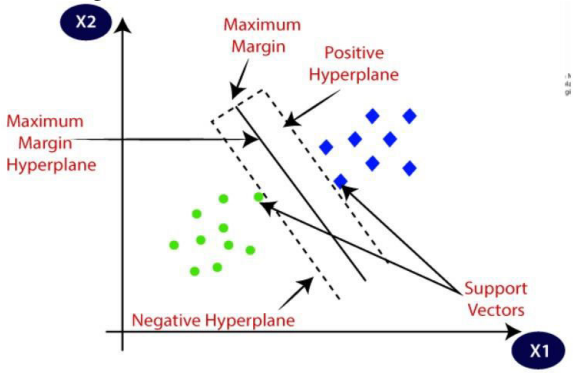

Fig 2: Support Vector machine

IV. RESULTS

The machine learning algorithms described in this work can complement medical professionals, but should not be used instead of them in selecting diagnostic pathways for classifiers. By using these technologies, many of the healthcare constraints associated with inaccurate diagnoses, missing information, cost and time can be reduced. By increasing awareness of risk factors and diagnostic variables, the application of ML approaches can help reduce the overall public health burden of liver disease worldwide. More importantly, in chronic liver disease, ML can reduce kidney-related mortality, transplantation and/or hospitalization by detecting kidney disease in its early or latent stages. Early recognition improves the prognosis because treatment can be started before the disease progresses further. Invasive tests such as biopsy would also be less likely in this scenario. Although this study focused on ML training for signs of hepatitis and chronic kidney disease, it is possible that the techniques can be applied to distinguish between other types of kidney disease and healthy individuals. All of the above techniques can be used in other medical disciplines to enable AI/ML-assisted diagnosis. Decision tree algorithms such as Random Forest, SVM, KNN, Logistic Regression were used in the study to predict liver disease at an early stage. These algorithms provide different results based on execution time, Kappa statistic, precision, mean absolute error, precision and recall.

The table lists the performance metrics used in the comparison. Using decision trees to predict kidney disease helps manage people's health. However, we continue to collect the latest data from multiple locations around the world to diagnose kidney disease. Liver and heart related diseases become more common with time. With the continuous development of technology, they will only increase in the future. Although people today are more health conscious and join yoga classes, dance classes; still a sedentary lifestyle and luxury that is constantly being introduced and developed; the problem has been going on for a long time. So in such a situation our project is very useful for the society. With the data set used in this project, we achieved 78 accuracy for the SVM model, and although this accuracy can be difficult to achieve with very large data sets, the results of this project clearly show that we can predict kidney diseases with precision. of 90% or more.

V. CONCLUSION

The effectiveness of these methods was compared and evaluated. Analysis shows that Decision Stump outperforms other algorithms with 72.67% accuracy. The table lists the performance metrics used in the comparison. Using decision trees to predict kidney disease helps manage people's health. However, we continue to collect the latest data from

multiple locations around the world to diagnose kidney disease. Liver and heart related diseases become more common with time. With the continuous development of technology, they will only increase in the future. Although people today are more health conscious and join yoga classes, dance classes; still a sedentary lifestyle and luxury that is constantly being introduced and developed; the problem has been going on for a long time. So in such a situation our project is very useful for the society. With the data set used in this project, we achieved 78 accuracy for the SVM model, and although this accuracy can be difficult to achieve with very large data sets, the results of this project clearly show that we can predict kidney diseases with precision. of 90% or more.

REFERENCES

[1]. M. P. N. M. Wickramasinghe, D. M. Perera and K. A. D. C. P. Kahandawaarachchi, "Dietary prediction for patients with Chronic Kidney Disease (CKD) by considering blood potassium level using machine learning algorithms," 2017 IEEE Life Sciences Conference (LSC), Sydney, NSW, 2018, pp. 300-303.
[2]. H. A. Wibawa, I. Malik and N. Bahtiar, "Evaluation of Kernel-Based Extreme Learning Machine Performance for Prediction of Chronic Kidney Disease," 2018 2nd International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, 2018, pp. 1-4
[3]. U. N. Dulhare and M. Ayesha, "Extraction of action rules for chronic kidney disease using Naïve bayes classifier," 2020 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2020, pp. 1-5.
[4]. H. Zhang, C. Hung, W. C. Chu, P. Chiu and C. Y. Tang, "Chronic Kidney Disease Survival Prediction with Artificial Neural Networks," 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 2019, pp. 1351-1356
[5]. J. Aljaaf et al., "Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics," 2019 IEEE Congress on Evolutionary Computation (CEC), Rio de Janeiro, 2018, pp. 1-9.

BIOGRAPHY

Mr.P..Manivannan,As sociateProfessor, DepartmentofElectronicsandCommunicationEngineering,Adhiya maanCollegeofEngineering, AnnaUniversity.

Prabhu.M, DepartmentofElectronicsandCommunicationEngineering,Adhi yamaanCollegeofEngineering, AnnaUniversity.

Kannan.B, DepartmentofElectronicsandCommunicationEngineering,Adhi yamaanCollegeofEngineering, AnnaUniversity.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 11, Special Issue 1, February 2023 ||

Harishkumar.S,
DepartmentofElectronicsandCommunicationEngineering,AdhiyamaanCollegeofEngineering,
AnnaUniversity.

Loka prabhu .M,
DepartmentofElectronicsandCommunicationEngineering,AdhiyamaanCollegeofEngineering,
AnnaUniversity.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details