IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CKD PREDICTION SYSTEM USING MACHINE LEARNING

[1]Deepak K N, [2]Adhwaidh P S, [3]Akshay P D, [4]Athira K S, [5]Jisna Jayan

1Assistant Professor, Department of Computer Science and Engineering, [2] Student, Department of Computer Science and Engineering, [3] Student, Department of Computer Science and Engineering

[4]Student, Department of Computer Science and Engineering, 5Student, Department of Computer Science and Engineering

Universal Engineering College, Thrissur, Kerala

Abstract: Chronic kidney disease (CKD) is a global health issue that causes a high rate of morbidity and mortality, as well as the onset of additional diseases. Because there are no clear symptoms in the early stages of CKD, people frequently miss it. Early identification of CKD allows patients to obtain timely treatment to slow the disease's progression. Due to their rapid and precise recognition capabilities, machine learning models can successfully assist doctors in achieving this goal. We propose a machine learning framework for diagnosing CKD in this paper. The CKD data set was taken from kaggle, which has a substantial number of missing values. We employ multiple machine learning methods such as DT, SVM, and DNN to analyze data from CKD patients with 21 characteristics and 400 records. The dataset is preprocessed by filling in missing data and normalizing it. To increase accuracy and save training time, the most relevant features from the dataset are chosen. Image processing and letter recognition are used to automatically input the attributes.

Index Terms **– Chronic kidney disease, Support vector machine, Decision tree, Deep neural network, OCR.**

I. INTRODUCTION

Kidney infection is the major health problem that affects people all around the world. As life is so costly in our country, and people's daily and monthly incomes are so low, patients don't even go to hospitals and clinics; they can simply take painkillers, and most hospitals and clinics aren't prepared to manage kidney problems early. "Tolerating that risk Factors are recognised early, unusual kidney injury and reliable renal issue can be ruined," according to the WHO, "and, assuming kidney infection is checked early, disabling of kidney breaking point can be pushed back or prevented by simple intervention." Chronic kidney diseases include problems that harm the kidneys and reduce their ability to keep us healthy. If somehow the kidney problem worsen, waste can build up to dangerous quantities in our blood, causing problems like hypertension, weak bones, weakness, poor nutrition, and nerve damage. Moreover, kidney disease increases the risk of heart and vascular disease. Diabetes, hypertension, coronary artery disease, lupus, weakness, germs and egg whites in the urine, difficulties taking specific medications, salt and potassium deficiency in the blood, and familial ancestry are just a few of the causes of CKD. Early exposure and clinical diagnosis can prevent the progression of ongoing kidney disease. If it continues to worsen, it could lead to renal failure and the need for dialysis. It can also lead to the need for a kidney transplant to keep you alive. In this paper, we will use a machine learning prediction method to create a model that can detect the early stages and progression of chronic kidney disease, assisting wellness experts in their decision-making to avoid the CKD problem.

II. LITERATURE SURVEY

Jiongming Qin and Lin Chen [1] presented a machine learning algorithm for diagnosing CKD in their paper. KNN imputation was used to fill in the missing values in a dataset obtained from the University of California Irvine (UCI) machine learning repository, which selects several complete samples with the most similar measurement to process the missing data for each incomplete dataset. They use six machine learning algorithms to process the missing data for each incomplete dataset, including logistic regression, random forest, support vectors machine, K- nearest neighbour, naive bays classification, and naive bays classification. Among these machine learning models, random forest has the best diagnostic accuracy, with 99.75 %, and they analyse the misjudgment generated by the established model to propose an integrated model that combines logistic regression and random forest with perception, with an average accuracy of 99.83 %.

Zixian Wang and Jae Won Chung [2] presented a machine learning technique using data from the UCI machine learning dataset warehouse. For 400 chronic kidney patients, CKD is diagnosed using apriori association approaches using a 10-fold cross-validation test, and the results are compared across a range of classification algorithms including ZeroR, OneR, naive Bayes, J48, and IBK (K-nearest Neighbour). The dataset was pre-processed by filling in missing data and normalising it. To enhance accuracy and save training time, they selected the most relevant feature from the dataset. Based on Apriori, the detection accuracy for CKD is 99 %.

Shanila Yunus Yashfi and Md Ashikul Islam [3] collects data from CKD patients and proposes a technique for predicting CKD risk. This system should have used 456 patient data from the online UCI machine learning repository as well as a real-time dataset from the Khulna City Medical College. This system was developed using the Python high-level interpreted programming language. This system is trained with the data using 10-fold CV and applied random forest and ANN. Random forest method achieves 97.12 %, whereas ANN achieves 94.5 %.

A lab dataset of 361 patients with chronic renal disease was used by A.Radya and Ayman S. Anwar [4]. It uses the PNN, SVM, and MLP algorithms to calculate the duration of chronic renal disease. It was compared the results of methods like Probabilistic Neural Networks (PNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Radial Basis Function (RBF). The probabilistic brain organisation calculation appears to be the best performing calculation that clinicians can employ to avoid diagnostic and treatment errors, according to this study.

Pramila Arulanthu and Eswaran Perumal [5] proposed a feature selection method which is used to reduce the attributes and select only the most important ones. JRip, SMO, Naive Bayes, and IBK are the four classifiers used to classify data. By comparing the outcomes of the reduced attribute dataset with the original dataset using these four classifiers, this system can predict the correct and best classifiers. Classification is the most significant element of the process, which is done utilising data mining techniques based on machine learning. Data instance group membership can be predicted using the classification.

Himanshu Kriplani and Sudipta Roy [6] proposed a deep neural network-based method that accurately predicts the presence or absence of chronic renal disease with a % accuracy rate. This model, which is implemented using the cross-validation technique to keep the model safe from overfitting, produces better results when compared to other available algorithms. This automatic chronic kidney disease treatment slows the course of kidney damage, but it requires early detection of chronic kidney disease.

N V Ganapathi Raju and K Prasanna Lakshmi [7] presented a system to detect chronic renal disease by imposing several categorization algorithms on the patient's medical record. This research work is largely focused on determining the best suited classification algorithm that can be used for the diagnosis of CKD based on the classification report and performance criteria. Support Vector Machine, Random Forest, XGBoost, Logistic Regression, Neural Networks, and Naive Bayes Classifier are among the methods studied empirically. The experimental results show that Random Forest and XGBoost outperform other classification algorithms, achieving 99.29 percent accuracy.

Mrs Prasuna Kotturu and Mr VVS Sasank [8] proposed a technique that uses machine learning algorithms to detect chronic kidney disease (CKD). The supervised machine learning techniques used in this paper are random forest, support vector machine (SVM), linear regression (LR), decision tree, and naive Bayes classifier. The random forest strategy is the best suitable technique for detection after comparing the results produced utilising different methods. Random forest is used to achieve 99.3 % detection accuracy in this paper.

Naganna Chetty and Kunwar Singh Vaisla [9] uses classification models which has been built with different classification algorithms, Wrappersubset attribute evaluator and best first search method to predict and classify the CKD and non CKD patients. These models were tested on a recently gathered CKD dataset obtained from the University of California at Irvine's repository. The models did a better job of distinguishing between CKD and non-CKD cases. The outcomes of various models are compared. Classifiers performed better on the smaller dataset than on the original dataset, according to the comparison.

K.Shankar and P. Manickam [10] proposes a methodology using inspired optimization model and learning procedure to classify CKD. To choose suitable aspects of kidney data for the classification process, the suggested method utilizes the Ant Lion Optimization (ALO) strategy. The CKD data is then sorted using a Deep Neural Network based on chosen features (DNN). Performance comparisons reveal that this suggested model obtains greater classification accuracy, precision, F-measure, and sensitivity measures when compared to previous data mining classifiers.

Gunarathne W.H.S.D and Perera K.D.M [11] proposes a methodology to predict the patient's status of CKD or non CKD. Machine learning classification techniques were employed to predict the value. Different classification methods have been used to create classification models that will predict the patient's CKD and non-CKD status. These models were tested using a recently obtained CKD dataset of 400 data records and 25 characteristics, which was downloaded from the UCI repository. The outcomes of various models are compared. The model with the Multiclass Decision forest algorithm scored best in the comparison, with an accuracy of 99.1 % for the reduced dataset with 14 attributes.

S.Ramya and Dr. N.Radha [12] proposed a method for determining kidney function failure by using a classification algorithm on test results obtained from the patient's medical report. The purpose of this study is to use classification algorithms to speed up diagnosis and improve accuracy. The proposed work focus on classifying the severity of various stages of chronic kidney disease. Different algorithms are used in the experiment, including Back Propagation Neural Networks, Radial Basis Functions, and Random Forest. The testing findings reveal that the Radial basis function algorithm outperforms the other classification algorithms, with an accuracy rate of 85.3 %.

S.Dilli Arasu and Dr. R.Thirumalaiselvi [13] examine several data mining approaches in the medical area, as well as some of the algorithms that are used to predict kidney disease. The results of this study show that depending on the tools and techniques used, results for different stages of kidney disease diagnosis may vary. When appropriate techniques are used, data mining provides better outcomes in disease diagnosis.

As a result, data mining is an important field for healthcare prediction. Asif Salekin and John Stankovic [14] use novel approach to detect CKD using machine learning algorithm. They get results from a dataset with 400 records and 25 variables that indicate whether a patient has CKD or not. To get results, they use k-nearest neighbours, random forest, and neural networks. They use a wrapper method to reduce the number of features, which allows them to detect CKD with high accuracy.

When developing a neural network classifier for medical decision making on chronic kidney disease, Pinar Yildirim proposed this experiment [15] to analyze the effects of class imbalance in training data. Data mining and decision systems are two applications where neural networks are commonly used. Back propagation networks are a popular type of neural networks that can be trained to recognise different patterns. The importance of these networks was examined, and a comparative evaluation of some sampling methods based on multilayer perceptron with various learning rate values for the prediction of chronic kidney disease was performed. This study shows that sampling algorithms can increase classification algorithm performance, and learning rate is a critical parameter that has a major impact on multilayer perceptron.

Sahil Sharma and Vinod Sharma [16] assessed 12 different classification algorithm on dataset which having 400 records and 24 attributes. They calculated the accuracy of prediction results by comparing their calculated results to actual results. Accuracy, sensitivity, precision, and specificity were utilised as assessment metrics. The decision tree technique provides accuracy of up to 98.6%, sensitivity of 0.9720, precision of 1 and specificity of 1.

Tabassum S and Mamatha Bai B G [17] proposed a system using on big data in healthcare which has been developed by using data mining techniques. Data mining techniques have been used to analyse and forecast Chronic Kidney Disease (CKD). The clustering technique Expectation Maximization [EM] is used to put people who have comparable characteristics into one group. The classification algorithms Artificial Neural Network [ANN] and C4.5 are used to forecast disease.

Dr. S. Vijayarani and Mr. S. Dhayanand [18] proposed a method for predicting kidney diseases using Nave Bayes and Support Vector Machine classification methods. This study was mainly concerned with finding the optimum classification algorithm based on classification accuracy and execution time. Based on the results, it can be stated that the SVM improves classification performance and produces accurate results, making it the superior classifier when compared to the Nave Bayes classifier method. Perhaps the Nave Bayes classifier classifies the data with the least amount of processing time.

Sirage Zeynu and Shruti Patil [19] used K-Nearest Neighbor, J48, Artificial Neural Network, Naïve Bayes, and Support Vector Machine classification techniques to diagnose Chronic Kidney Disease. Create two key models to predict chronic kidney disease. The feature selection method and the ensemble model, to be specific. The wrapper subset evaluator with the best first engine and the Info gain attributes evaluator with ranker search engine were used to construct chronic kidney disease prediction. The results showed that the K-nearest neighbour classifier with Wrapper Sub set Evaluator and Best first search engine feature selection method has 99 % accuracy, J48 with Info Gain Attribute Evaluator and ranker search engine has 98.75 % accuracy, Artificial Neural Network with Wrapper Sub set Evaluator and Best first search engine has 99.5 percent accuracy, Nave Bayes with Wrapper Sub set Evaluator and Best first search engine has 99 % accuracy, Support Vect with Wrapper Sub The second model-building method uses a voting algorithm to combine five heterogeneous classifiers into an ensemble model. By comparing the proposed ensemble model to the base classifier, the effectiveness of the proposed ensemble model was determined. The proposed ensemble model was found to be 99 % accurate in the experiments.

Ani R1 and Greeshma Sasi1 [20] developed a clinical decision support system using machine learning techniques. The classification techniques such as neural network-based back propagation (BPN), probability-based Naive Bayes, LDA classifier, lazy learner K Nearest Neighbor (KNN), tree-based decision tree, and Random subspace classification algorithms are examined first in this study. On a dataset collected from the UCI repository with 25 attributes and 400 instances, the accuracy of each method identified is 81.5 %, 78 %, 76 %, 90 %, 93 %, and 94 %, respectively. The algorithm which produced the best results was used to develop the Clinical Decision Support System based on the data acquired.

III. RESEARCH AND METHODOLOGY

3.1 Data set

The CKD dataset was collected from 400 patients from the Kaggle. Dataset comprises 21 features divided into 16 numeric features and 5 categorical features, in addition to the class features, such as “ckd” and “notckd” for classification. Features include age, blood pressure, specific gravity, albumin, sugar, red blood cells, pus cell, pus cell clumps, bacteria, blood glucose random, blood urea, serum creatinine, sodium, potassium, hemoglobin, packed cell volume, white blood cell count, red blood cell count, hypertension, diabetes mellitus, coronary artery disease, appetite, pedal edema, and anemia. The diagnostic class contains two values: ckd and notckd. All features contained missing values except for the diagnostic feature.

3.2 System Design

In this study, several methods, procedures, and techniques have been used to build machine learning models. In this section, first we discussed dataset. Second, we made clear the data preprocessing methods. Finally, we give details on the machine learning algorithms for developing a CKD prediction model and explain the performance evaluation methods. The classifiers were first established by different machine learning algorithms to diagnose the data samples. The final result is chosen from the three individual outputs with the most votes.

Based on features, the proposed system is divided into three modules. The registration module is the first. It is used for login and signup purposes. When users sign up or register, their required data is stored in a database table so that when they try to login again, the username and password are cross-checked, and if they match, the user can sign in and gain access to the application. The next module is the scanning module, where the lab report is uploaded using OCR and Tesseract and returns its corresponding key value pairs. The classifier module is the final module. Support vector machines, decision trees, and deep neural networks are the three models used for classification.

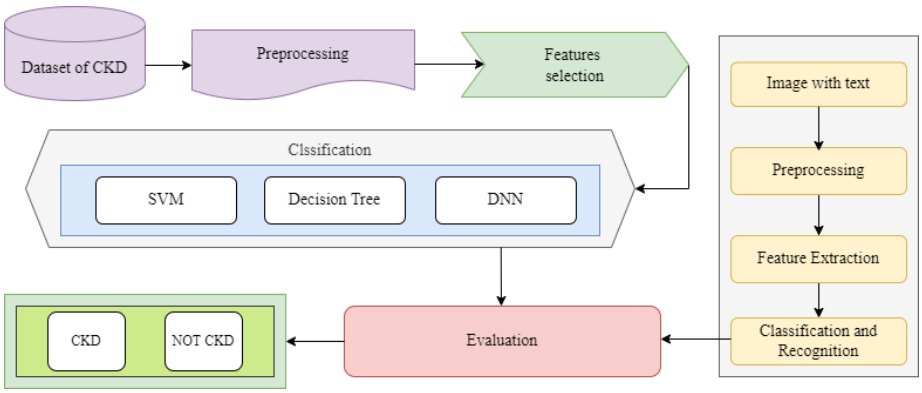

Fig (a) Block diagram of CKD prediction system

3.3 Support Vector Machine

This algorithm is the highest well-known and important supervised machine-learning algorithm that is works on classification and regression problems however mostly used for classification problem. SVM used kernel function to separate labeled data. One of the advantages associated with using kernels in SVM is that SVM applies kernel dentitions to non-vector inputs and Kernels can be defined based on a combination of different data types. In SVM algorithm data classify into two data points in which hyper-plane lies between two branches. SVM creates a discrete hyper plane in the signifier space of the training data and compounds are classified based on the side of the hyper plane located.

SVM has been utilized in several fields, including bioinformatics and handwritten recognition. SVM is also used in other applications, including medical diagnosis, weather prediction, stock market analysis, and image processing. Similar to all other machine learning techniques, SVM is a computational algorithm that learns from experience and examples to allocate labels to objects.

3.4 Decision Tree

Decision tree algorithm belongs to supervised machine learning algorithm and used to solve both classification and regression problems, but mostly used for classification problems. Decision Tree algorithm resolve the classification problem by converting the dataset into a tree representation through sorting them by feature values. In decision tree every node indicates features in an instance to be classified and every leaf node indicates a class label the instance belongs to.

3.5 Deep Neural Network

Deep neural network (DNN) is a type of machine learning algorithms in which the input and output layers are separated by numerous layers. The main aim is to imitate the data handling of the brain. DNN shave more than one secret layer arranged between the input and output layers. One of the key reasons deep learning is more remarkable than classical machine learning is that it makes adaptable arrangements. A model can be worked with a solitary layer of neurons, and adding layers allows the computer to make an ever increasing number of explicit elements that lead to a more perplexing last result.

3.6 Image processing

Image processing is a strategy for performing a specific procedure on an image in order to improve it or extract useful data from it. It is a type of sign processing in which the input is a picture and the output could be a picture or qualities/features related to that picture. Tesseract is an open source tool for recognizing text in images using Optical Character Recognition (OCR) method of extracting text from images. Deep learning is applied in Tesseract OCR. It is complex and the results are high accuracy. Tesseract library is used to perform OCR on images and the output is stored in a text file. Diagnosis tests include different types of information, such as symptoms and medical tests. Doctor's conclusion of medical treatment rest on diagnosis tests which makes the accuracy of diagnosis is important in medical care. Providentially, the attributes of the diagnosis tests can be measured for a given disease condition.

IV. RESULT AND DISCUSSION

To predict the performance of the prediction model we need to use accuracy, confusion matrix: precision, recall. We calculate true positive (TP), True Negative (TN), false Positive (FP) and false Negative (FN).

True Positive (TP) is a condition that prediction model correctly predicted the positive class which means the prediction and the actual class are positive. True Negative (TN) is a condition that prediction model correctly predict the negative class which means both the prediction class and actual class are negative. False Positive (FP) is a condition that prediction model gives the wrong prediction of the negative class which means prediction class is positive and actual class is negative. False Negative (FN) is a condition that the prediction models wrongly predict the positive class which means prediction class is negative and actual class is positive.

4.1 Accuracy

Accuracy means the algorithm capability to predict the class of the dataset correctly. It determines of how close or near the predicted value is to the actual value. Accuracy computes using the following equation:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

4.2 Precision

Precision measure the true values correctly predicted from the total predicted values in the actual class. Precision quantifies the ability of the classifiers to not label a negative example as positive. The equation to calculate precision is:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

4.3 Recall

Recall determine the rate of the total positive values that are correctly predicted. Recall answers what percentage of actual Positives is correctly classified. The equation to calculate Recall calculated is:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

Confusion Matrix is one of the tools for evaluating the behavior a binary classifier. For better visualization of results, we have used heat maps for each model as shown in figures (b) to (d).

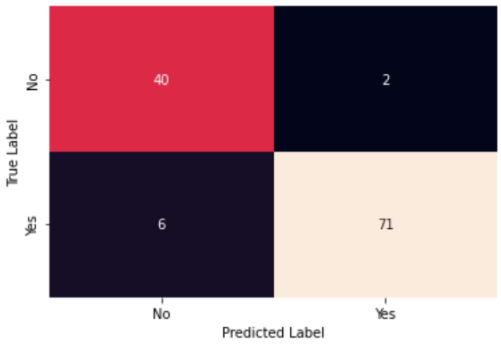

Fig (b) SVM Classifier Result

As visible in figure 8.1 SVM is tested on 120 sample among 40 sample classified as non-ckd, 71 sample as ckd and remain 9 classify as false.

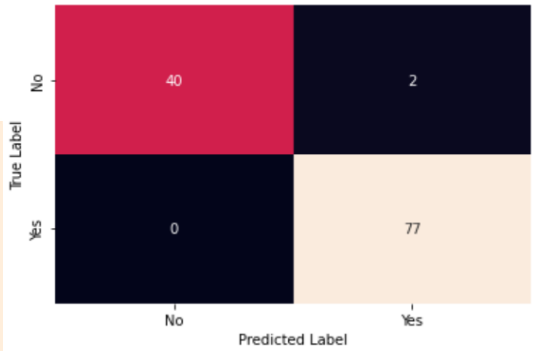

Fig (c) DT Classifier Result

As visible in figure 8.2 DT is tested on 120 sample among 40 sample classified as non-ckd, 77 sample as ckd and remain 3 classify as false.

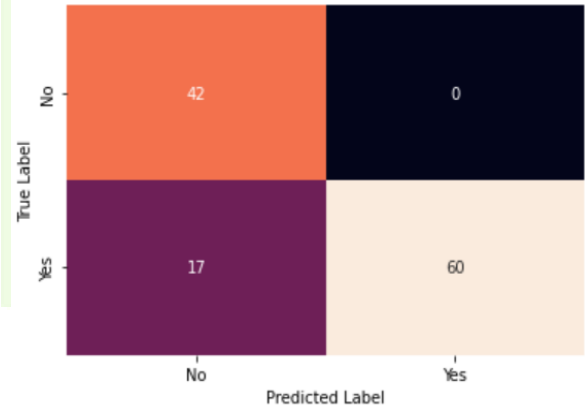

Fig (c) DNN Classifier Result

As visible in figure 8.1 DNN is tested on 120 sample among 42 sample classified as non-ckd, 60 sample as ckd and remain 18 classify as false. The results indicate that DT has achieved high performance. DNN has not performed well in comparison, other machine learning algorithms.

Table 1 Predictive Accuracies of Classification Algorithms on Data set

Classifier	Accuracy %
SVM	93.28
DT	98.32
DNN	92.46
Overall accuracy	99.15

V. CONCLUSION

This paper describes a clinical area application that uses machine learning algorithms to assist clinical professionals in anticipating CKD based on CKD parameters. Support Vector Machine, Decision Tree, and Deep Neural Network are three machine learning algorithms that have been considered for predicting chronic kidney disease. The dataset is obtained from kaggle. The dataset contains 400 records and 26 features. From 26 features, we only use 21 of them. The features are automatically entered using image processing and letter recognition. This system will identify chronic kidney disease in early stages. This assists specialist with diagnosing and recommending the treatment at the beginning phase. Patients will be able to learn about their health condition at an earlier stage and follow the necessary diet and prescriptions to improve the progression of this condition.

REFERENCES

[1]. Y. Liu, C. Liu, J. Qin, L. Chen, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease," IEEE Access, vol. 8, pp. 20991–21002, 2020, doi: 10.1109/ACCESS.2019.2963053

[2]. Z. Wang, J. Won Chung, X. Jiang, Y. Cui, M. Wang, and A. Zheng, "Machine Learning-Based Prediction System For Chronic Kidney Disease Using Associative Classification Technique," Int. J. Eng. Technol., vol. 7, no. 4.36, p. 1161, 2018, doi: 10.14419/ijet.v7i4.36.25377.

[3]. S. Y. Yashfi et al., "Risk Prediction of Chronic Kidney Disease Using Machine Learning Algorithms," 2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020, 2020, doi: 10.1109/ICCCNT49239.2020.9225548.

[4]. A. S. Anwar and E. H. A. Rady , "Prediction of kidney disease stages using data mining algorithms," Informatics Med. Unlocked, vol. 15, no. March, p. 100178, 2019, doi: 10.1016/j.imu.2019.100178.

[5]. E. Perumal , P. Arulanthu , "Predicting the Chronic Kidney Disease using Various Classifiers," 4th Int. Conf. Electr. Electron. Commun. Comput. Technol. Optim. Tech. ICEECCOT 2019, pp. 70–75, 2019, doi: 10.1109/ICEECCOT46775.2019.9114653

[6]. H. Kriplani, B. Patel, and S. Roy, "Prediction of chronic kidney diseases using deep artificial neural network technique," vol. 31. Springer International Publishing, 2019.

[7]. N. V. Ganapathi Raju, K. Prasanna Lakshmi, K. G. Praharshitha, and C. Likhitha, "Prediction of chronic kidney disease (CKD) using Data Science," 2019 Int. Conf. Intell. Comput. Control Syst. ICCS 2019, no. Iciccs, pp. 642–647, 2019, doi: 10.1109/ICCS45141.2019.9065309.

[8]. 9.P. Kotturu, V. V. S. Sasank, G. Supriya, C. S. Manoj, and M. V. Maheshwarredy, "Prediction of chronic kidney disease using machine learning techniques," Int. J. Adv. Sci. Technol., vol. 28, no. 16, pp. 1436–1443, 2019, doi: 10.17148/IJARCCE.2018.71021.

[9]. S. D. Sudarsan,N. Chetty, K. S. Vaisla, "Role of attributes selection in classification of Chronic Kidney Disease patients," pp. 1–6, 2016, doi: 10.1109/cccs.2015.7374193.

[10]. K. Shankar, G. Devika, P. Manickam, M. Ilayaraja, "Optimal Feature Selection for Chronic Kidney Disease Classification using Deep Learning Classifier," 2018 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2018, pp. 1–5, 2018, doi: 10.1109/ICCIC.2018.8782340

[11]. GunarathneW.H.S.D,Perera K.D.M, Kahandawaarachchi K.A.D.C.P, "Performance Evaluation on Machine Learning Classification Techniques for Disease Classification and Forecasting through Data Analytics for Chronic Kidney Disease (CKD)",2017 IEEE 17th International Conference on Bioinformatics and Bioengineering

[12]. S.Ramya, Dr. N.Radha, "Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms," Proc. International Journal of Innovative Research in Computer and Communication Engineering,Vol. 4, Issue 1, January 2016.

[13]. S.DilliArasu and Dr. R.Thirumalaiselvi, "Review of Chronic Kidney Disease based on Data Mining Techniques",International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 23 (2017) pp. 13498-13505

[14]. Asif Salekin, John Stankovic, "Detection of Chronic Kidney Disease and Selecting Important Predictive Attributes," Proc. IEEE International Conference on Healthcare Informatics (ICHI), IEEE, Oct. 2016, doi:10.1109/ICHI.2016.36.

[15]. Pinar Yildirim, "Chronic Kidney Disease Prediction on Imbalanced Data by Multilayer Perceptron: Chronic Kidney Disease Prediction," Proc. 41st IEEE International Conference on Computer Software and Applications (COMPSAC), IEEE, Jul. 2017, doi: 10.1109/COMPSAC.2017.84

[16]. Sahil Sharma, Vinod Sharma, Atul Sharma, "Performance Based Evaluation of Various Machine Learning Classification Techniques for Chronic Kidney Disease Diagnosis," July18, 2016.

[17]. Tabassum S, [2] Mamatha Bai BG, [3] Jharna Majumdar," Analysis and Prediction of Chronic Kidney Disease using Data Mining Techniques," International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 4, Issue 9, September 2017.

[18]. Dr.S. Vijayarani1, Mr.S.Dhayanand2," DATA MINING CLASSIFICATION ALGORITHMS FOR KIDNEY DISEASE PREDICTION," International Journal on Cybernetics & Informatics (IJCI) Vol. 4, No. 4, August 2015.

[19]. S. Zeynu and S. Patil, "Prediction of Chronic Kidney Disease using Data Mining Feature Selection and Ensemble Method," International Journal of Data Mining in Genomics and Proteomics, 9(1), pp. 1-9, 2018.

[20]. R. Ani, S. Greeshma, S. Resmi, and O. Deepa, "Decision Support System for Diagnosis and Prediction of Chronic Renal Failure Using Random SubSpace Classification," International. Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, pp. 1-6, 2016.