International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

EARLY PREDICTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING TECHNIQUES: A REVIEW

T Suguna , L Sumathi

Abstract

Chronic Kidney Disease (CKD) is a major public health problem, affecting more than 10% of the world's population. Early diagnosis and treatment of CKD is essential to reduce the burden of the disease and improve patient outcomes. Machine Learning has been attracting many researchers, and it has been successfully applied in many fields such as banking, e-commerce, and healthcare etc. In recent years, machine learning (ML) techniques have been used to develop predictive models for CKD. This survey paper reviews the current state of the art of ML based CKD prediction, focusing on the most relevant papers published in the last decade. The survey paper discusses the data sets used for training and testing, challenges, and limitations of existing CKD prediction models, and provides recommendations for future research. Additionally, this survey paper provides a comprehensive comparison of the performance of the various machine learning algorithms and techniques used for CKD prediction.

Keywords:

Chronic Kidney Disease, Deep Belief Network, Machine Learning, Early Prediction, SVM

1. INTRODUCTION

Chronic kidney disease (CKD) is a long-term, progressive condition that affects the kidneys' ability to filter waste and excess fluids from the blood. It is characterized by a gradual loss of kidney function over time, which can lead to a build-up of toxins in the body. Some common causes of CKD include diabetes, high blood pressure, glomerulonephritis, and polycystic kidney disease. Symptoms may not appear until the disease is in its advanced stages, but they can include fatigue, difficulty concentrating, decreased appetite, swelling in the feet and ankles, and changes in urine output. Diagnosis of CKD is typically done through blood tests and urine tests, which can measure kidney function and detect the presence of protein or blood in the urine. Treatment options vary depending on the underlying cause of the disease and the stage of CKD. Lifestyle changes such as controlling blood pressure, managing diabetes, and quitting smoking can help slow the progression of the disease. Medications, dialysis, and kidney transplant may also be necessary for advanced cases of CKD. This global public health issue affects 10% of the world's population [1], [2]. In China, the prevalence of CKD is 10.8% [3], but in the United States, it ranges from 10% to 15% [4]. Among the overall adult population of Mexico 14.7% [5] of adult are affected by CKD. This illness is characterized by a gradual decline in renal function that ultimately results in a total loss of renal function. Early on, CKD does not have noticeable symptoms. Because of this, the illness might not be discovered until the kidney has lost around 25% of its functionality [6]. Moreover, CKD has a significant worldwide influence on the human body and high morbidity and mortality [7]. It has been linked to the development of cardiovascular disease and pathologic syndrome [8], [9].

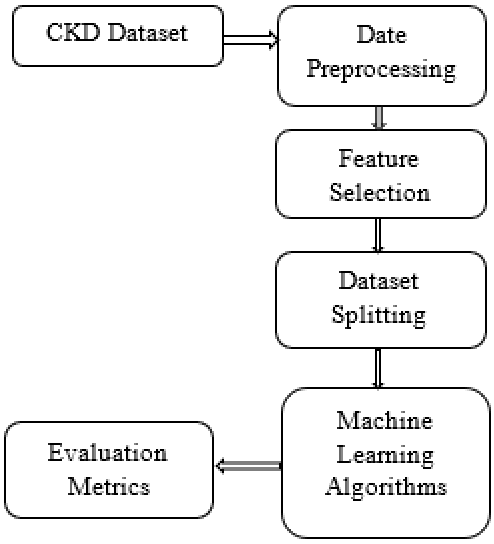

Figure.1 Architecture of CKD Prediction

As a result, predicting and diagnosing CKD in its early stages is critical, as it may allow patients to receive timely treatment to slow the disease's progression. Early detection and diagnosis of CKD can be challenging due to the lack of specific symptoms in its early stages. Machine learning algorithms have shown great potential in identifying the early stages of CKD, and the authors of this paper have investigated the use of machine learning in predicting CKD. The overall process of prediction of CKD using machine learning is shown in the figure1.

2. DESCRIPTION OF THE DATA SET

Table 1. Description of the dataset

S.No	Attribute	Description	Type	Value
1.	age	Age	numeric	Years
2.	bp	Blood Pressure	numeric	mm/Hg
3.	sg	Specific Gravity	numeric	1.005, 1.010, 1.015, 1.020, 1.025
4.	al	Albumin	numeric	0, 1, 2, 3, 4, 5
5.	su	Sugar	numeric	0, 1, 2, 3, 4, 5
6.	rbc	Red Blood Cells	nominal	normal, abnormal
7.	pc	Pus Cell	nominal	normal, abnormal
8.	pcc	Pus Cell Clumps	nominal	present, notpresent
9.	ba	Bacteria	nominal	present, notpresent
10.	bgr	Blood Glucose Random	numeric	mgs/dl
11.	bu	Blood Urea	numeric	mgs/dl
12.	sc	Serum Creatinine	numeric	mgs/dl
13.	sod	Sodium	numeric	mEq/l
14.	pot	potassium	numeric	mEq/l
15.	hemo	Hemoglobin	numeric	Gms
16.	pcv	Packed Cell Volume	numeric	-
17.	wbcc	White Blood Cell Count	numeric	cells/cumm
18.	rbcc	Red Blood Cell Count	numeric	millions/cmm
19.	htn	Hypertension	numeric	yes, no
20.	dm	Diabetes Mellitus	numeric	yes, no
21.	cad	Coronary Artery Disease	nominal	yes, no
22.	appet	Appetite	nominal	good, poor
23.	pe	pedal Edema	nominal	yes, no
24.	ane	Anemia	nominal	yes, no
25.	Class	Class	nominal	ckd, notckd

UCI Machine learning Dataset [23] consists of 25 attributes including 11 numeric and 14 nominal data. Table1 shows the attribute and its type, value, and description.

3. RELATED WORK

3.1. DIETARY PREDICTION FOR PATIENTS WITH CHRONIC KIDNEY DISEASE (CKD) BY CONSIDERING BLOOD POTASSIUM LEVEL USING MACHINE LEARNING ALGORITHMS [10]

Chronic kidney disease (CKD) is a long-term medical condition that affects the kidneys and can lead to a variety of health complications. To manage CKD, it is important to monitor the patient's dietary intake and ensure that they are consuming the right amount of nutrients. This study aimed to develop a machine learning algorithm to predict dietary intake for patients with CKD based on their blood potassium level. The study used a dataset of 400 patients with CKD, which included their dietary intake and blood potassium levels. The dataset was divided into a 70 % training set and a 30% testing set. Models used in this study are Multiclass Decision Jungle, Decision Forest, Neural Network and Logistic Regression. Among the four models Multiclass Decision Forest provides 99.17% accuracy, which is best compared to other algorithms.

There are six stages in the CRISP-DM such as Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment. After the deployment, the model is used to get the different diet plans for CKD patients according to their blood potassium level by using the predicted potassium level. The accuracy of different machine learning algorithm used in this work is shown in table2. The main advantage of this study is that it demonstrates the potential of machine learning algorithms to accurately predict dietary intake for patients with CKD. This could be used to help health care providers better to manage the dietary needs of their patients. The main limitation of this study is that it only used a dataset of 400 patients with CKD. This may not be representative of the entire population of patients with CKD, and further research is needed to validate the results of this study. Additionally, the study did not consider other factors that may affect dietary intake, such as lifestyle and medical history.

Table 2. Accuracy

Algorithm	Overall Accuracy
Multiclass DF	99.17%
Multiclass Decision Jungle	97.50%
Multiclass LR	89.17%
Multiclass NN	82.50%

3.2 PREDICT CHRONIC KIDNEY DISEASE USING DATA MINING ALGORITHMS IN HADOOP [11]

This research is focuses on developing a predictive model to identify patients with chronic kidney disease (CKD) using data mining This research is focuses on developing a predictive model to identify patients with chronic kidney disease (CKD) using data mining algorithms in Hadoop. The paper utilizes a CKD dataset with clinical parameters of the patients, such as age, gender, body mass index, creatinine level, and blood pressure. The paper utilizes the Hadoop distributed computing platform to process the dataset. The data mining algorithms used in the study include KNN and support vector machines. The algorithms are used to analyze the data and create a predictive model to identify patients with CKD.

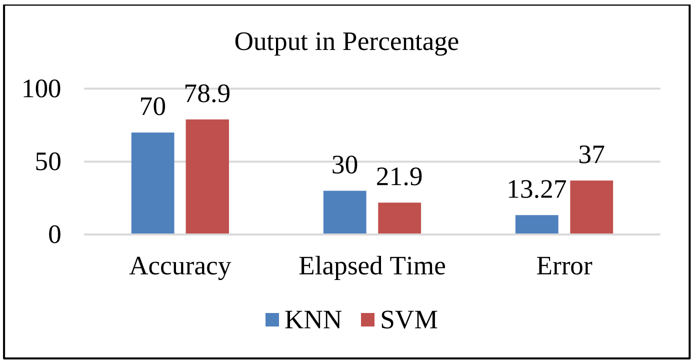

Figure 2. Results of KNN and SVM

The results of the study showed that the support vector machine algorithm had the highest accuracy in correctly identifying patients with CKD. The use of Hadoop distributed computing platform allows for faster and more efficient processing of the dataset. The use of data mining algorithms allows for the development of a predictive model that can accurately identify patients with CKD. The study was limited to a dataset of limited patients, which may not be representative of the general population. Additionally, the predictive model developed in the study may not be applicable to other populations.

3.3 EARLY PREDICTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING SUPPORTED BY PREDICTIVE ANALYTICS [12]

This approach creates a predictive model for checking CKD using predictive analysis. This paper utilizes the best collection of parameters to feed into the predictive model to do the predictive analysis. There are 24 parameters in the dataset. Three major categories comprise the distribution of parameters. The dataset comprises of 37.5% Healthy data and 62.5% CKD data. Binary numerical parameters are utilised to indicate normal instances and anomalies. Binary 0 and binary 1 are used in this context. Outliers or extreme values are crucial to this work. Some of the parameters required to build the predictive model includes sodium, potassium, RBC and WBC count, RBC in urine. This is used to examine the relevance of parameters and association between the parameters those are extracted from Blood Haematology, Urine Test, Other general information and clinical factors. There are seven parameters are considered for optimal subset such as Blood glucose, Haemoglobin, specific gravity, albumin, pus cells, hypertension and creatinine. Regression Tree (RPART), SVM, LR, Multi-Layer Perceptron (MLP) are used for model building. Performance metrics used for the result analysis are Sensitivity(TPR), Specificity(TNR), Precision(PPV), Accuracy, AUC,F1score. The result is as shown in table 2.

Table 3. Comparative Results

ML Methods	TPR	TNR	PPV	ACC	AUC	F1 Score
RPART	0.9339	1	1	0.956	0.982	0.965
SVM	0.9892	0.8955	0.9292	0.950	0.973	0.958
LR	0.9897	0.9677	0.9797	0.9797	0.994	0.981
MLP	0.9897	0.9677	0.9797	0.9797	0.995	0.981

3.4 A MACHINE LEARNING METHODOLOGY FOR DIAGNOSING CHRONIC KIDNEY DISEASE [13]

This paper presents a machine learning methodology for diagnosing CKD using a combination of data mining, feature selection, and classification techniques. The methodology was tested on a dataset of patients with CKD and compared with traditional methods for CKD diagnosis. The machine learning methodology used in this paper consists of three steps: data mining, feature selection, and classification. The data mining step involves extracting relevant features from the dataset.

Feature selection is then used to select the most informative features for CKD diagnosis. Finally, a classification algorithm is used to classify the patients into CKD or non-CKD classes. The classification algorithms used in this work are Logistic Regression, SVM, KNN, Naive Bayes, and DNN. Performance metrics used to evaluate and compare the performance of the machine learning methodology include accuracy, sensitivity, specificity, and F-measure. Accuracy is the proportion of correct predictions made by the model, while sensitivity and specificity are measures of how well the model can distinguish between CKD and non-CKD patients. The F-measure is a weighted average of the precision and recall of the model. Methodology of using KNN Imputation and an integrated model is feasible to help classify CKD diagnosis. The accuracy is 99.83%. Results of Integrated model for the various imputation(K) values are shown in the figure.

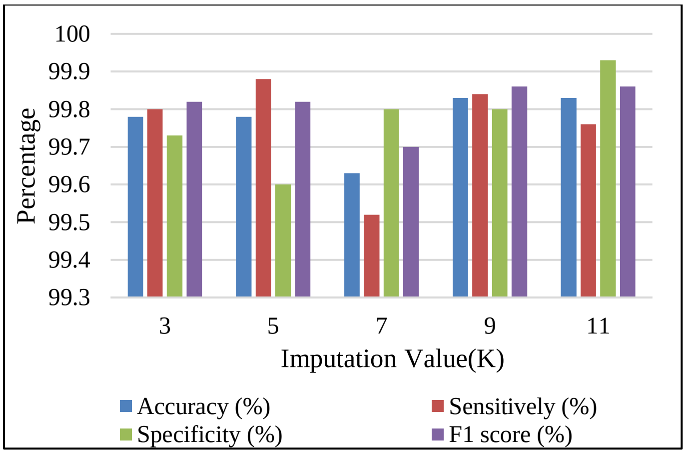

Figure 3. Comparative results for different K value

The main advantage of the machine learning methodology is that it can accurately diagnose CKD with higher accuracy than traditional methods. The methodology is also able to select the most informative features from the dataset, which can be used to improve the accuracy of the model. One of the main disadvantages of the machine learning methodology is that it requires a large amount of data for training and testing. Additionally, the model may be prone to overfitting if the dataset is not properly balanced.

3.5 PREDICTION OF CHRONIC KIDNEY DISEASE (CKD) USING DATA SCIENCE [14]

Data-driven model to predict chronic kidney disease (CKD) using data science is developed in this work. The model uses a combination of machine learning algorithms and statistical techniques to analyze patient data and identify individuals at risk of CKD. The model is trained on large datasets of patient information and uses the patient's demographic, clinical, and laboratory data to make predictions. Statistical techniques such as feature selection and feature engineering are used to identify relevant features and reduce the dimensionality of the data. The model uses a variety of metrics to evaluate the performance of the model, including accuracy, precision, recall, and F1 score.

In this paper CKD was predicted using SVM, Random Forest, XGBoost, LR, Naïve Bayes algorithms. There are five stages in this work namely Acquire, Data Pre-processing, Data Exploration, Feature Selection, and Model Selection. In Acquire stage, the data set is obtained from UCI repository which contains 400 patients records. In Data pre-processing the CKD and NOCKD labels are replaced by binary values as 1 and 0. And the missing values are replaced by median values respectively. Data Exploration is done in the beginning of the data analysis to comprehend what is present in the dataset and to know the traits of the data. Here univariant and bi-variant analysis are used. Univariant analysis is the graphical representation of the dataset attributes. These are of 3 types such as Gplot, Distribution curve and PCA (Principle Component Analysis) used to reduce the dimensionality of the dataset. In Bi-variant analysis the correlation between the attributes is calculated. The fourth step is Feature Selection. Using univariant feature selection, the top 5 features are selected. The features are pcv, hemo, sc, rbcc, and sg. The last stage is Model Selection. Here the Five algorithms namely SVM, LR,RF, XGBoost, NB are used to build model to predict CKD. The result of method is shown in the table 3.

Table 4. Precision and FI-Score

Classifier	Precision	F1-Score
LR	1	0.98
SVM	0.98	0.98
XGB	0.99	0.99
NB	1	0.98
RF	1	0.99

The model can accurately predict CKD with high accuracy and precision. Additionally, the model is generalizable and robust, and can be applied to a variety of datasets. The model is limited in its ability to identify individuals at risk of CKD in the early stages of the disease. Additionally, the model is limited in its ability to identify individuals with CKD who are not yet diagnosed.

3.6 EARLY DETECTION AND PREVENTION OF CHRONIC KIDNEY DISEASE [15]

In this study, the elements that can increase a patient's likelihood of having CKD are extrapolated using a Data Mining method and Boruta analysis. This analysis includes historical, statistical, and medical information. The dataset was gathered from a UCI source and includes information on 400 samples from the southern region of India, whose ages ranged from 2 to 90 years. No single factor in the dataset can definitively determine whether a person has chronic kidney disease. However, CKD can be identified by characteristics such as hypertension and blood pressure, which are some of the main causes of the condition, as well as by straightforward tests such as blood pressure, urine albumin, and serum creatinine. These elements are therefore regarded as risk factors. One positive test does not always indicate the presence of CKD; rather, a combination of these tests is frequently needed. When monitoring kidney function, a variety of metrics can be used. These include measuring levels of creatinine and urea in the blood, as well as measuring glomerular filtration rate (GFR).

3.7 CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING [16]

This research paper [] explores the potential of machine learning for predicting chronic kidney disease (CKD). The authors used a dataset of 400 patients with CKD and 400 healthy patients to build a machine learning model. The model was trained using a variety of algorithms, including Ant Colony Optimization and support vector machines. There are 3 stages in the proposed work. The are Pre-Processing, Feature Selection and Classification. In Pre-Processing, Missing values are replaced by the estimated mean value. Data objects are labelled by the number of features, the label encoding is done by converting each unique attribute value to an integer. Ant Colony Optimization is the technique used for the feature selection. This algorithm will change the small number of features in subsets which are selected by choosing the best ants. Support Vector Machine is used for classification to predict the CKD. The main aim is to predict the patients with CKD using minimum number of attributes while maintaining high accuracy. Accuracy of 96% was obtained in this method. The comparative result is shown in the table

Table 5. Comparative Results for ML algorithms

	Precision	Recall	F1-Score	Support
CKD	1.00	0.940	0.97	83
NO CKD	0.88	1.00	0.94	37
Macro-avg	0.94	0.97	0.95	120
Weighted Avg	0.96	0.96	0.96	120

3.8 PERFORMANCE ANALYSIS OF MACHINE LEARNING CLASSIFIER FOR PREDICTING CHRONIC KIDNEY DISEASE [17]

This is a study that uses machine learning algorithms to predict chronic kidney disease (CKD). The purpose of this study was to compare the performance of various machine learning algorithms for the prediction of CKD. The dataset used for this study was collected from the UCI Machine Learning Repository. The dataset consists of 400 instances and 25 attributes, including clinical parameters such as age, blood pressure, and serum creatinine. Methods Used The study used several machine learning algorithms to predict CKD, including decision tree and random forest. The performance of these algorithms was evaluated using several metrics, including accuracy, precision, recall, and F-measure. The algorithms were compared using 10-fold cross-validation. The proposed work is to predict Chronic Kidney Disease using Decision Tree, Random Forest and Logistic Regression. Different process of the work are Dataset preparation, Data mining and Preprocessing, Feature Selection , performance evalution and Classification. Performance metrics are evaluated using TP(True positive), TN(True Negative), FP(False Positive), and FN(False Negative) values.

Table 6. Comparative Results

Classifier	Accuracy	Precision	Recall
Decision Tree	98.48%	100.0	97.62
Random Forest	94.16%	95.12	96.29
LR	99.24%	98.82	100.0

3.9 EARLY PREDICTION OF CHRONIC KIDNEY DISEASE USING DEEP BELIEF NETWORK [18]

This research paper proposes a deep belief network (DBN) model for predicting chronic kidney disease (CKD). The DBN model is based on a combination of a deep learning algorithm and a Bayesian network. The deep learning algorithm is used to learn the features of the dataset and the Bayesian network is used to infer the probability of CKD from the features. The model is trained on a dataset of CKD patients and then tested on a separate dataset. The model is evaluated using a variety of metrics, including accuracy, precision, recall, and F1 score. The results of the model are compared to traditional machine learning models and other deep learning models. The results of the DBN model proposed show that the model can accurately predict CKD with an accuracy of 98.52%, and an F1 score of 87.5%. The results of the model are compared to traditional machine learning models and other deep learning models, and the DBN model is shown to outperform all of them. Results of the metrics are shown in the table 7.

Table 7. Results of Training Data and Testing data

Metrics	Training data	Testing data
Accuracy	98.52	98.51
Precision	89.95	86.62
Recall	87.70	87.50
F-Measure	87.70	87.50
RMSE	0.484552	0.482874
Mean Absolute Error	0.3766666	0.37

3.9 DIAGNOSIS OF CHRONIC KIDNEY DISEASE USING EFFECTIVE CLASSIFICATION ALGORITHMS AND RECURSIVE FEATURE ELIMINATION TECHNIQUES [19]

This paper focuses on the use of effective classification algorithms and recursive feature elimination techniques to diagnose chronic kidney disease (CKD). The paper explores the use of machine learning techniques to identify and classify CKD patients. The dataset used for the study was obtained from the

UCI Machine Learning Repository. The dataset contains 24 clinical parameters (features) of 400 patients. The paper uses several machine learning algorithms to classify CKD patients. These algorithms include logistic regression, support vector machines (SVM), Naive Bayes, k-nearest neighbors (KNN), and decision tree (DT). The authors also use recursive feature elimination (RFE) to reduce the number of features in the dataset. RFE is used to identify the most important features that are most relevant to the diagnosis of CKD. The authors use several metrics to evaluate the performance of the machine learning algorithms. These include accuracy, precision, recall, F1 score, and the area under the receiver operating characteristic curve (AUC). The results of the study shown in the table 8.

Table 8. Comparison of Precision, Recall, Accuracy and F1 Score

Classifiers	SVM	KNN	Decision tree	Random forest
Accuracy %	96.67	98.33	99.17	100.00
Precision %	92.00	100.00	100.00	100.00
Recall %	94.74	97.37	98.68	100.00
F1-score%	97.30	98.67	99.34	100.00

3.10 CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING ENSEMBLE ALGORITHM [20]

The goal of this paper is to forecast chronic kidney disease using machine learning ensemble techniques. The UCI repository's data set was utilised to create the model [8]. 400 patients' records with 25 attributes, including the class, are included in the collection. The dataset includes of information gathered from blood and urine tests as well as some general details like age and hunger. Out of the 400 individuals, 150 were healthy, and 250 had CKD as their diagnosis. Training set and testing set are created from the dataset. With the use of several machine learning ensemble techniques, the training set is used to create the model. Each ensemble classifier's hyper parameters are adjusted to obtain the optimal parameters for the best model of patient-specific chronic kidney disease prediction. The testing dataset is then used with the trained model. According to each model's performance in terms of accuracy, sensitivity, specificity, precision, F-score, ROCAUC, and Mathew Correlation Coefficient, the model is evaluated, and the result is shown in table 9.

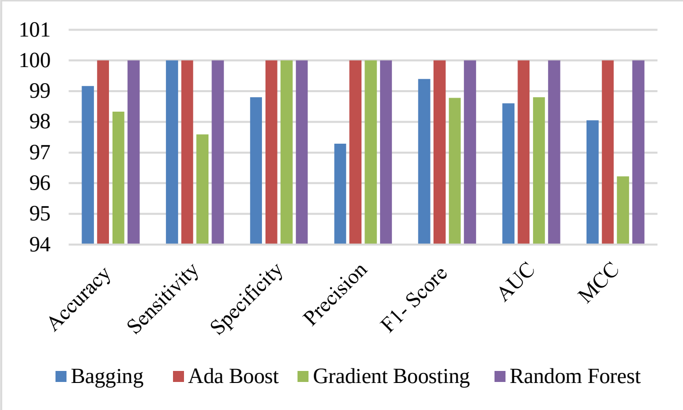

Figure 4. Results of various ensemble algorithms

3.11 PREDICTION OF CHRONIC KIDNEY DISEASE - A MACHINE LEARNING PERSPECTIVE [21]

The paper discusses the various machine learning algorithms such as Artificial Neural Network, C5.O, Chi-Square, Automatic Interaction Detector, Logistic Regression, Linear Support Vector Machine with penalty L1 and penalty L2, and Random tree that can be used to build a predictive model for CKD, and the performance of these models when evaluated using various metrics. The work also discusses the potential benefits of using machine learning for predicting CKD and the challenges associated with it. The results are computed based on full features, correlation based feature selection, wrapper method feature selection, least absolute shrinkage and selection operator regression, synthetic minority over sampling technique with full feature and synthetic minority over sampling technique with least feature. The major steps involved in this work are Dataset Pre-processing, Feature Selection, Classifier Application, SMOTE, and Analysing the classifier. The paper evaluates the performance of the predictive models using a variety of metrics, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve. The figure shows the performance of the various Classifier Models with different feature selection method.

3.12 A COMPARATIVE STUDY, PREDICTION AND DEVELOPMENT OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING ON PATIENTS CLINICAL RECORDS [22]

Due to use of clinical variables as a training dataset the model can predict CKD patients at an early stage. Preprocessing the data after collection is essential to getting it ready for modelling. Predictive mean matching (PMM) has been used to handle missing values as one of the crucial steps that make up data preparation. Data clustering has been applied to the dataset in order to identify the different groups within the data collection. Using both the original features and the features chosen through the feature selection procedure, the performance of the classifiers' evaluated. The 7 most crucial features out of the 25 features were chosen using the XGBoost-based feature selection method. Two different datasets namely the main dataset and the XGBoost dataset were made for training purposes. After feature selection, creating classification models necessitates dividing the data into training and testing sets, with 80% of the data going into training and the remaining 20% going into testing. When building a model, the preparation set is used to evolve the model, while the experiment set is used to legalize the model. Later, used diferent ML plans containing NN, RF, SVM, RT, and BTM to expect useful patterns. XGBoost dataset achieved best accuracy of 100% by using SVM method. The comparative result of different parameters are shown in figure 5 and Table 9.

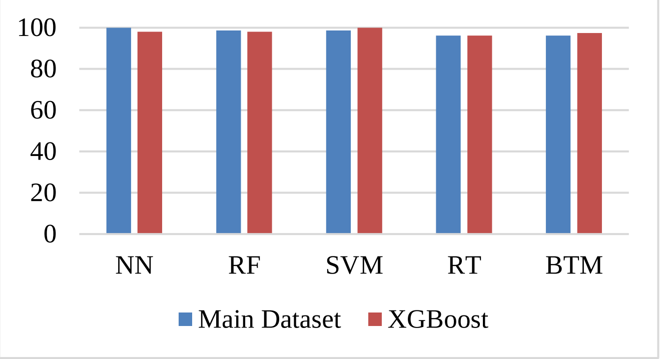

Figure 5. Results of Accuracy

Table 9. Comparison of main dataset and XG Boost dataset

Algorithm	Main Dataset (25 attributes)			XG Boost dataset (7 attributes)		
	Sensitivity	specificity	Kappa	Sensitivity	specificity	Kappa
NN	100	100	100	98.00	96.67	94.67
RF	98.04	100	97.32	98.04	100	97.32
SVM	98.04	100	97.32	100	100	100
RT	94.34	100	91.84	94.34	100	91.84
BTM	94.34	100	91.84	96.15	100	94.59

4. CONCLUSION

This survey paper has examined the current state of Chronic Kidney Disease (CKD) prediction using machine learning techniques. By analyzing various standard research papers, provided an overview of the different machine learning algorithms and techniques used for CKD prediction, as well as the data sets used for training and testing. It also discussed the challenges and limitations of existing CKD prediction models and provided recommendations for future research. Additionally, this survey paper has provided a comprehensive comparison of the performance of the various machine learning algorithms and techniques used for CKD prediction. The results of this survey paper suggest that the use of machine learning algorithms and techniques can provide an effective and reliable way to predict CKD. Future research should focus on improving existing models and developing new models that can accurately predict CKD.

5. REFERENCES

[1] Z. Chen et al., “Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers,” Chemometr. Intell. Lab., vol. 153, pp. 140-145, Apr. 2016.

[2] A. Subasi, E. Alickovic, J. Kevric, “Diagnosis of chronic kidney disease by using random forest,” in Proc. Int. Conf. Medical and Biological Engineering, Mar. 2017, pp. 589-594.

[3] L. Zhang et al., “Prevalence of chronic kidney disease in china: a crosssectional survey,” Lancet, vol. 379, pp. 815-822, Aug. 2012.

[4] A. Singh et al., “Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration,” J. Biomed. Inform., vol. 53, pp. 220-228, Feb. 2015.

[5] A. M. Cueto-Manzano et al., “Prevalence of chronic kidney disease in an adult population,” Arch. Med. Res., vol. 45, no. 6, pp. 507-513, Aug. 2014.

[6] H. Polat, H.D. Mehr, 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.

[7] C. Barbieri et al., “A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease patients undergoing dialysis,” Comput. Biol. Med., vol. 61, pp. 56-61, Jun. 2015.

[8] V. Papademetriou et al., “Chronic kidney disease, basal insulin glargine, and health outcomes in people with dysglycemia: The origin study,” Am. J. Med., vol. 130, no. 12, Dec. 2017.

[9] N. R. Hill et al., “Global prevalence of chronic kidney disease - A systematic review and meta-analysis,” Plos One, vol. 11, no. 7, Jul. 2016.

[10] 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, Australia, 2017, pp. 300-303, doi: 10.1109/LSC.2017.8268202.

[11] G. Kaur and A. Sharma, "Predict chronic kidney disease using data mining algorithms in hadoop," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 973-979, doi: 10.1109/ICICI.2017.8365283.

[12] A. J. Aljaaf et al., "Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics," 2018 IEEE Congress on Evolutionary Computation (CEC), Rio de Janeiro, Brazil, 2018, pp. 1-9, doi: 10.1109/CEC.2018.8477876.

[13] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng and B. Chen, "A Machine Learning Methodology for Diagnosing Chronic Kidney Disease," in IEEE Access, vol. 8, pp. 20991-21002, 2020, doi: 10.1109/ACCESS.2019.2963053.

[14] N. V. Ganapathi Raju, K. Prasanna Lakshmi, K. G. Praharshitha and C. Likhitha, "Prediction of chronic kidney disease (CKD) using Data Science," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 642-647, doi: 10.1109/ICCS45141.2019.9065309.

[15] M. Desai, "Early Detection and Prevention of Chronic Kidney Disease," 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, 2019, pp. 1-5, doi: 10.1109/ICCUBEA47591.2019.9128424.

[16] Reshma S, Salma Shaji, S R Ajina, Vishnu Priya S R, Janisha A, "Chronic Kidney Disease Prediction using Machine Learning," in International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 07, July-2020.

[17] Rahul Gupta , Nidhi Koli , Niharika Mahor , N Tejashri, Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease, in 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020

[18] S. M. M. Elkholy, A. Rezk and A. A. E. F. Saleh, "Early Prediction of Chronic Kidney Disease Using Deep Belief Network," in IEEE Access, vol. 9, pp. 135542-135549, 2021, doi: 10.1109/ACCESS.2021.3114306.

[19] E Jadhav, and Mohammed Y. Alzahrani , Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques, Hindawi Journal of Healthcare Engineering Volume 2021, Article ID 1004767, 10 pages https://doi.org/10.1155/2021/1004767

[20] Nikhila, "Chronic Kidney Disease Prediction using Machine Learning Ensemble Algorithm," 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2021, pp. 476-480, doi: 10.1109/ICCCIS51004.2021.9397144.

[21] P. Chittora et al., "Prediction of Chronic Kidney Disease - A Machine Learning Perspective," in IEEE Access, vol. 9, pp. 17312-17334, 2021, doi: 10.1109/ACCESS.2021.3053763.

[22] Hassan, M.M., Hassan, M.M., Mollick, S. et al. A Comparative Study, Prediction and Development of Chronic Kidney Disease Using Machine Learning on Patients Clinical Records. Hum-Cent Intell Syst (2023). https://doi.org/10.1007/s44230-023-00017-3

[23] https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease