FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 12, December, 2025, pp 72 – 80
DOI: https://doi.org/10.33003/fjs-2025-0912-4025

PERFORMANCE EVALUATION OF MACHINE LEARNING ALGORITHM USING CHRONIC KIDNEY DISEASE (CKD) DATASET

***[1]Abdul Malik Aliyu, [1]Kabiru Abdul Mumin, [1]Abdullahi Bashar Abubakar, [1]Abubakar Suleiman Abba and [2]Abubakar Ibrahim**

[1]Department of Computer Science, College of Science and Technology, Umaru Ali Shinkafi Polytechnic, Sokoto, Nigeria.
[2]Department of Mathematics and statistics, College of Science and Technology, Umaru Ali Shinkafi Polytechnic, Sokoto, Nigeria.

*Corresponding authors' email: akay2kay@yahoo.com

ABSTRACT

Machine learning (ML) algorithms enable computers to recognize patterns and make predictions or decisions from data, rather than relying on explicit programming. This paper presents a predictive model for the early detection of CKD through ML. The study uses five years of Electronic Health Record (EHR) data from a diverse patient group. The dataset contains demographics, clinical history, lab results, medication information, and diagnostic codes. The research starts with 25 variables, in addition to the class property, and then reduces this to 15 by using Principal Component Analysis (PCA). This aims to reduce the number of parameters to find the best subset for identifying CKD. The research uses common ML algorithms—Support Vector Machines (SVM), Random Forest, and Logistic Regression—and assesses their ability to detect CKD early. When comparing the classification algorithms, Random Forest (RF) had the best accuracy, at 81.2967%.

Keywords: CKD- Chronic Kidney Disease, EHR- Electronic Health Record, ML- Machine Learning, PCA- Principal Component Analysis, RF- Random Forest, SVM-Support Vector machine

INTRODUCTION

Machine learning, a subfield of artificial intelligence, helps predict outcomes. It helps extract information from reports and support decision-making through data analysis (Srivastava et al., 2020). Machine learning aids in creating predictive models to distinguish between defective and non-defective cases (Sharma, 2019). The purpose of machine learning is to construct computer systems that can adapt and learn from their experience. (Iliyas *et al*., 2020). Various ML algorithms, like Logistic Regression (LR), Support Vector Machine (SVM), Artificial Neural Network (ANN), K-Nearest Neighbor, Naïve Bayes classifier, and Random Forests, can predict the early stages of chronic kidney disease (CKD). The kidney is important for removing toxic and non-essential substances from the body through waste processing. It filters waste and excess fluids from the blood, which are then excreted (National Kidney Foundation Inc., 2024). Kidney failure is an important issue today. Chronic kidney disease involves the gradual loss of kidney function over time. Diabetes and high blood pressure are common causes. They affect millions globally, causing fatigue, drowsiness, itching, and pain (Islam, Majumder & Hussein, 2023). Smoking, diabetes, high blood pressure, heart disease, obesity, family history of renal problems, alcohol intake, age, race, sex, and drug use may raise the chance of kidney disease (Rahman et al., 2022). This study includes a comparison of different ML methods to get results. This research includes a comparison of ML models—Random Forest, Logistic Regression, and Support Vector Machine—using different methods to find the best-performing model. Electronic health record (EHR) data can be helpful for predicting CKD onset and progression. The spread of electronic health records (EHR) and machine learning (ML) offers chances for better disease understanding and risk prediction (Chen et al., 2019; Bernerjee, Chen, & Fatemifar, 2021). After data analysis, predictive models can be built using ML algorithms. These can then be checked using metrics like accuracy, sensitivity, and specificity to see how well they work. The goal of ML predictive modeling for early CKD detection is to check the performance of predictive models in terms of accuracy and sensitivity, and to find the best method for early detection of chronic kidney disease. This research aims to check how well ML algorithms perform in predicting chronic kidney disease by finding the key elements of the disease. We looked at EHR data over five years for patients with kidney disease, including those diagnosed with CKD and a control group without CKD.

Classification Algorithms

Three classification algorithms are used in this research. These algorithms were found to perform well in past studies. They are applied to the dataset to find the best one in terms of accuracy, sensitivity, specificity, and how long they take to run. The classification algorithms (SVM, RF, LR) are used to see which one predicts the early stage of CKD best and fastest.

Support Vector Machine

The Support Vector Machine (SVM) algorithm is used for classification and regression. It works well for classification problems where the goal is to divide data points into different categories. SVM finds the best hyperplane that separates the data into classes while maximizing the margin between the hyperplane and the nearest data points of each class (Rahman et al., 2022: Caraga et al., 2008). SVM is used to predict CKD by collecting and preparing a dataset with features and labels related to CKD, like age, blood pressure, and serum creatinine. The data is cleaned, normalized, and encoded for SVM. Equation (1) below shows the hyperplane that separates two classes.

$$D(\chi) = \omega o + \omega 1 a 1 + \omega 2 a 2 \qquad (1)$$

Figure (1) shows a visual representation of margins in support vector machine:

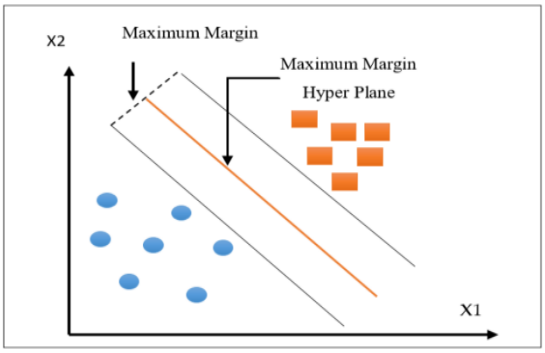


Figure 1: Support Vector Machine

Therefore, equation (2) shows the maximum-margin hyperplane

$\chi = b + \Sigma i\ \alpha i\ yi a(i) X\ a$ (2)

In this case i is the support vector, and y_i is the training instance a (i) class value. The learning algorithm determines the numeric value b and αi, respectively.

Random Forest Classifier

Random forest is a supervised machine learning algorithm that is used widely in Classification and Regression problems which builds decision trees on different samples and takes the majority vote of the data for classification and the average in the case of regression (khan *et al*., 2019). The random forest algorithm is based on ensemble learning, improving the model's performance, and solving complex problems by putting together several classifiers. The collected and preprocessed data being encoded to fit random forest are selected based on feature selection to reduce dimensionality and noise of data are trained and tested for random forest. The Random forest trained data is trained to suit a number of trees in its maximum debt and evaluated using accuracy, precision and recall matrices shown in Fig. (2).

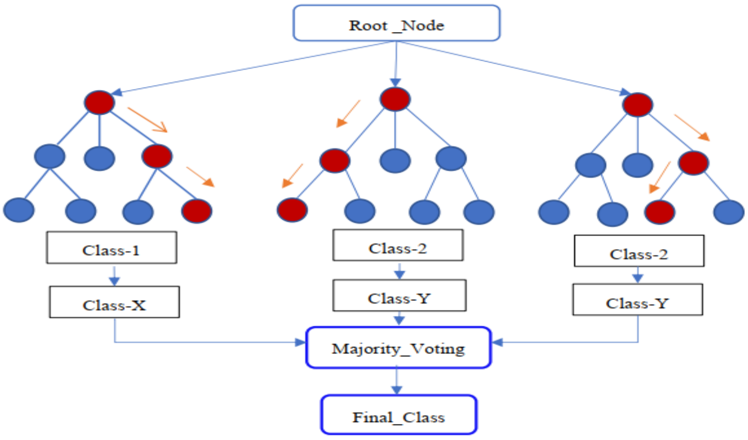


Figure 2: Random Forest

Logistic Regression

Logistic regression is a statistical model that uses a logistic feature to simulate a binary dependent variable in its simplest form. It is a widely used method for classification. This is a multiple regression variant where the expected result is binary instead of quantitative (Kempf-Leonard, 2004). Predicting CKD using Logistic regression, the collected and preprocessed data being encoded to fit logistic regression model are selected based on feature selection to reduce dimensionality to select the best features for the model, and are trained and tested for logistic regression.

Related Works

In recent years, the assimilation of Electronic Health Records (EHRs) has revolutionized healthcare data management, providing rich and comprehensive information on patient health.

This literature review focuses on the application of predictive modeling for early detection of CKD using machine learning techniques in which much work has been done with regards to the application on several machine learning techniques and algorithms in early detection of CKD. Among these are the most outperforming algorithms by (Rahman *et al*., 2022) where they predicted kidney disease by employing and comparing various machine learning algorithms including Logistic Regression, Naive Bayes, Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). They applied Principal Component Analysis (PCA) to reduce the dimensionality of the data and achieved 98% accuracy using the SVM technique that outperformed the rest of the models. Although, the study didn't include duration of execution of the models. (Gudeti *et al*., 2021) worked on three machine learning algorithms Logistic Regression, Support Vector Machine, and K-Nearest Neighbors where the results exemplified that the Support Vector Machine algorithm predicts Chronic Kidney Disease better with 0.9925187 accuracy than Logistic Regression having 0.7725 and K-Nearest Neighbors with 0.7875 accuracy level. it should be noted that the dataset they used was somewhat limited. Furthermore, duration of execution of these machine learning methods was not a factor in the investigation.

Ilyas *et al*., (2021) specifically used the Random Forest and J48 algorithms to obtain a sustainable and practicable model to detect the various stages of CKD with comprehensive medical accuracy. The study established and compared two algorithms J48 and random forest to predict the various stages of CKD (Stages 1-5). It is observed that the ratio of correctly classified instances and time taken by J48 is better than Random Forest. Hence, J48 is more accurate and efficient in terms of execution time because it provides results with better accuracy and less time than the Random forest.

Zahid & Mona, (2023) worked on developing prediction models for detecting and diagnosing CKD based on predominant features using machine learning techniques, to help reduce clinical expenses incurred by patients who are prescribed multiple identical tests. They employed K-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), and bagging, where KNN outperformed the other models in terms of accuracy, sensitivity, precision, specificity, F-measure and AUC score. The study had relatively small number of instances on the dataset, attributes such as GFR and eGFR which are also the main predictors for detecting CKD at the early stage.

MATERIALS AND METHODS

In this study, we employed three ML algorithms to predict kidney disease. Then, we looked at the features that led to high accuracy by getting information about those features in relation to kidney disease. Fig. 3 below Shows our research design.

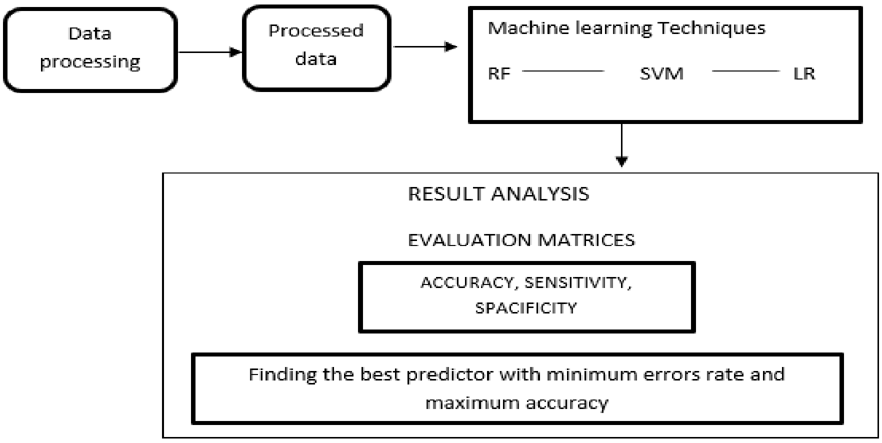

Data processing
Processed data
Machine learning Techniques
RF
SVM
LR
RESULT ANALYSIS
EVALUATION MATRICES
ACCURACY, SENSITIVITY, SPACIFICITY
Finding the best predictor with minimum errors rate and maximum accuracy

Figure 3: Research Design

To achieve this, we collected the CKD disease dataset and extracted the value for each feature. PCA was used to reduce the dataset's dimensions. The classification algorithms were applied to the new dataset to maximize accuracy. When the important value for each feature was calculated, the lowest important value for a few features was found. This showed that those features had little relation to our output. These features were combined into two principal components to reduce the dataset dimension.

Data Collection

To achieve this, we collected the CKD disease dataset and extracted the value for each feature. PCA was used to reduce the dataset's dimensions. The classification algorithms were applied to the new dataset to maximize accuracy. When the important value for each feature was calculated, the lowest important value for a few features was found. This showed that those features had little relation to our output. These features were combined into two principal components to reduce the dataset dimension. Table 1 below shows the distribution of predictor variables.

Table 1: Characteristics of Data Attributes and Predictor Variables

Feature	Specification	Value
AGE	AGE (IN YEARS)	0-90
AL	ALBUMEN	0-5
ANE	ANAEMIA	YES, NO
APPET	APPETIT	GOOD, POOR
BA	BACTERIA	PRESENT, NOTPRESENT
BGR	BLOOD GLUCOSE RANDOM	0-490
BP	BLOOD PRESSURE	0-180
BU	BLOOD UREA	0-391
CAD	CORONARY ARTERY DISEASE	YES, NO
CLASS	CLASS	CKD, NOTCKD
DM	DIABETES MELLITUS	YES, NO
HEMO	HAEMOGLOBIN	0-17.8
HTN	HYPERTENSION	YES, NO
PC	PUS CELL	NORMAL, ABNORMAL
PCC	PUS CELL CLUMPS	PRESENT, NOTPRESENT
PCV	PACKED CELL VOLUME	0-54

Feature	Specification	Value
PE	PEDAL EDEMA	YES, NO
POT	POTASSIUM	0-47
RBC	RED BLOOD CELLS	NORMAL, ABNORMAL
RBCC	RED BLOOD CELL COUNT	0-8u
SC	SERUM CREATININE	0-76u
SG	SPECIFIC GRAVITY	0-1.025u
SOD	SODIUM	0-163u
SU	SUGAR	0-5u
WC	WHITE BLOOD CELL COUNT	0-26,400u

Performance Evaluation Criteria

This work will employ a Predictive data analysis technique which utilizes historical and current facts to reach future predictions. It can also use data from a subject to predict the values of another subject. There are different predictive models; however, a simple model with more data can work better in general. Therefore, the prediction data set and also the determination of the measuring variables are important aspects to consider (MacGregor, 2013).

To understand the behavior of the classifiers, the following hypothesis are proposed:

True positive (TP) which will be the number of correctly predicted positive samples.

True negative (TN) will be the number of negative samples correctly predicted.

False negative (FN) will be the number of positive samples incorrectly predicted.

False positive (FP) will be the number of negative samples incorrectly predicted as positive.

Table (2) shows the performance criteria deployed to predict the machine learning algorithms in terms of accuracy, sensitivity and specificity.

Table 2: Performance Evaluation Criteria

Metric	Description	Formula
Accuracy	Number of correct predictions from all predictions	$\frac{TP+TN}{TP+FP+TN+FN}$
Sensitivity	Proportion of positive predictions that are correctly identified	$\frac{TP}{TP+FN}$
Specificity	Proportion of negative predictions that are correctly identified	$\frac{TN}{FP+TN}$

RESULTS AND DISCUSSION

Processing Result

The pre-processing result of the dataset in this study contains 25 attributes which are nominal and numerical values. Figure 4 shows a visualized display of the attributes contained in the dataset used in this study.

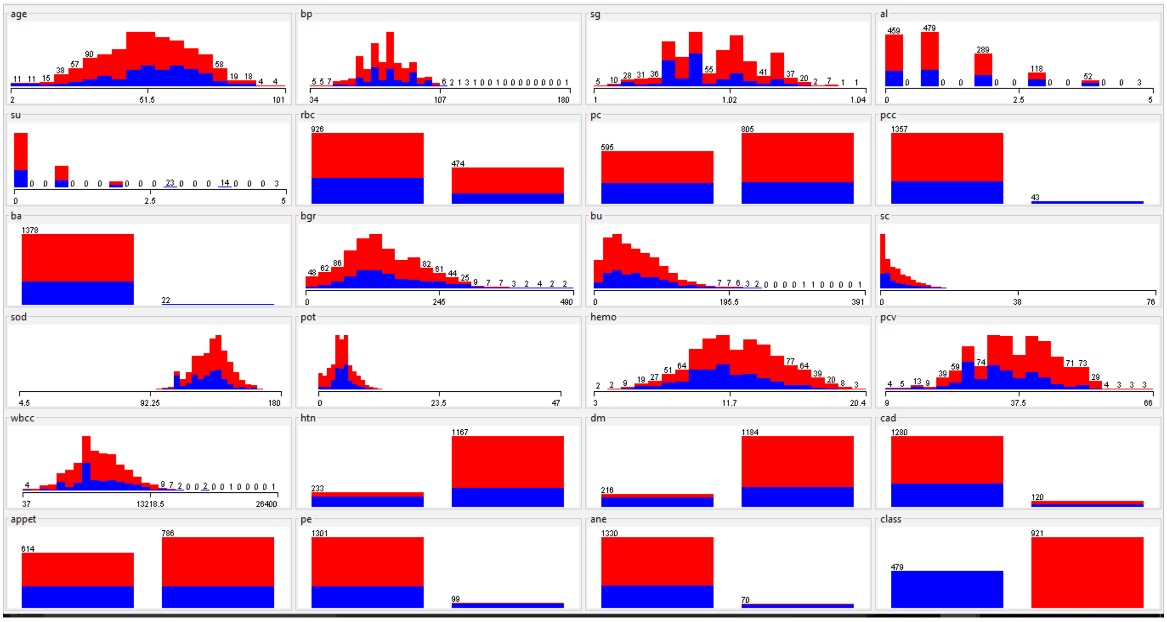
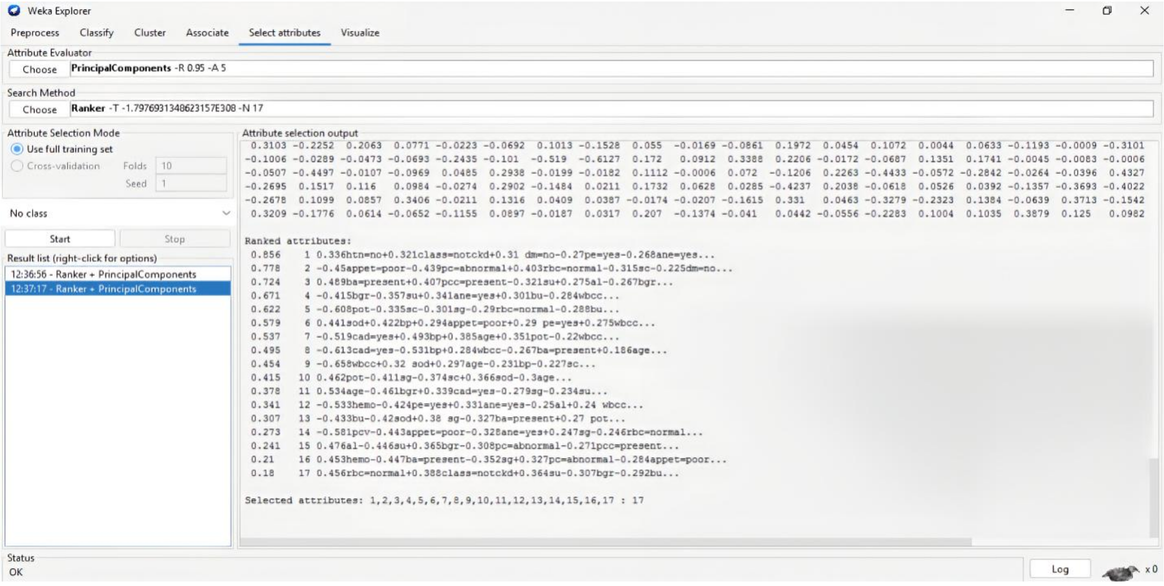


Figure 4: Visual Representation of Attributes

Performance Result of The Classifiers

This section entails the performance of the different models for predicting the early stage of chronic kidney disease using different performance metrics discussed in the methodology are demonstrated and also the experimental result of the algorithms experimented on CKD dataset is displayed. Three machine learning algorithms are utilized in this study which are: RF, SVM and LR. This study employs the PCA technique to reduce the dimensionality of the data for better prediction of the early stage of chronic kidney disease using PCA in line with the Ranker attributes by their individual attribute evaluation. The dataset comprising of 25 attributes was utilized, hence, Figure (5) shows the attributes calculated based on their importance value and found a low importance value for some features, which show that those features portray low value to our output. We chose features based on their importance value and utilized 10-fold cross validation for running the models.



Attribute Evaluator
Choose **PrincipalComponents** -R 0.95 -A 5

Search Method
Choose **Ranker** -T -1.7976931348623157E308 -N 17

Attribute Selection Mode
● Use full training set
○ Cross-validation Folds 10 Seed 1

No class

Result list (right-click for options)
12:36:56 - Ranker + PrincipalComponents
12:37:17 - Ranker + PrincipalComponents

Attribute selection output

```
 0.3103 -0.2252  0.2063  0.0771 -0.0223 -0.0692  0.1013 -0.1528  0.055  -0.0169 -0.0861  0.1972  0.0454  0.1072  0.0044  0.0633 -0.1193 -0.0009 -0.3101
-0.1006 -0.0289 -0.0473 -0.0693 -0.2435 -0.101  -0.519  -0.6127  0.172   0.0912  0.3388  0.2206 -0.0172 -0.0687  0.1351  0.1741 -0.0045 -0.0083 -0.0006
-0.0507 -0.4497 -0.0107 -0.0969  0.0485  0.2938 -0.0199 -0.0182  0.1112 -0.0006  0.072  -0.1206  0.2263 -0.4433 -0.0572 -0.2842 -0.0264 -0.0396  0.4327
-0.2695  0.1517  0.116   0.0984 -0.0274  0.2902 -0.1484  0.0211  0.1732  0.0628  0.0285 -0.4237  0.2038 -0.0618  0.0526  0.0392 -0.1357 -0.3693 -0.4022
-0.2678  0.1099  0.0857  0.3406 -0.0211  0.1316  0.0409  0.0387 -0.0174 -0.0207 -0.1615  0.331   0.0463 -0.3279 -0.2323  0.1384 -0.0639  0.3713 -0.1542
 0.3209 -0.1776  0.0614 -0.0652 -0.1155  0.0897 -0.0187  0.0317  0.207  -0.1374 -0.041   0.0442 -0.0556 -0.2283  0.1004  0.1035  0.3879  0.125   0.0982

Ranked attributes:
 0.856    1 0.336htn=no+0.321class=notckd+0.31 dm=no-0.27pe=yes-0.268ane=yes...
 0.778    2 -0.45appet=poor-0.439pc=abnormal+0.403rbc=normal-0.315sc-0.225dm=no...
 0.724    3 0.489ba=present+0.407pcc=present-0.321su+0.275al-0.267bgr...
 0.671    4 -0.415bgr-0.357su+0.341ane=yes+0.301bu-0.284wbcc...
 0.622    5 -0.608pot-0.335sc-0.301sg-0.29rbc=normal-0.288bu...
 0.579    6 0.441sod+0.422bp+0.294appet=poor+0.29 pe=yes+0.275wbcc...
 0.537    7 -0.519cad=yes+0.493bp+0.385age+0.351pot-0.22wbcc...
 0.495    8 -0.613cad=yes-0.531bp+0.284wbcc-0.267ba=present+0.186age...
 0.454    9 -0.658wbcc+0.32 sod+0.297age-0.231bp-0.227sc...
 0.415   10 0.462pot-0.411sg-0.374sc+0.366sod-0.3age...
 0.378   11 0.534age-0.461bgr+0.339cad=yes-0.279sg-0.234su...
 0.341   12 -0.533hemo-0.424pe=yes+0.331ane=yes-0.25al+0.24 wbcc...
 0.307   13 -0.433bu-0.42sod+0.38 sg-0.327ba=present+0.27 pot...
 0.273   14 -0.581pcv-0.443appet=poor-0.328ane=yes+0.247sg-0.246rbc=normal...
 0.241   15 0.476al-0.446su+0.365bgr-0.308pc=abnormal-0.271pcc=present...
 0.21    16 0.453hemo-0.447ba=present-0.352sg+0.327pc=abnormal-0.284appet=poor...
 0.18    17 0.456rbc=normal+0.388class=notckd+0.364su-0.307bgr-0.292bu...

Selected attributes: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17 : 17
```

Status
OK

Figure 5: PCA Selected Attributes

Figure 6. shows that the most impactful attributes of the dataset are htn, appet, ba, bgr, pot, sod, cad, wbcc, age, hemo, bu, pcv, al, bp, rbc and bp. The important values of these attributes are 0.856, 0.778, 0.724, 0.671, 0.622, 0.579, 0.495, 0.454, 0.378, 0.341, 0.307, 0.273, 0.241, 0.21 0.18, other features like pc, pcc, with attributes 0.07, 0.03 were found to be quite very low.

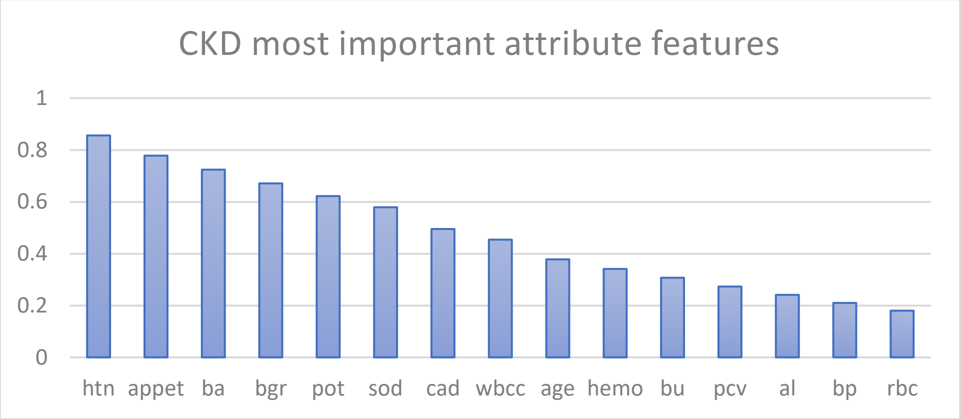
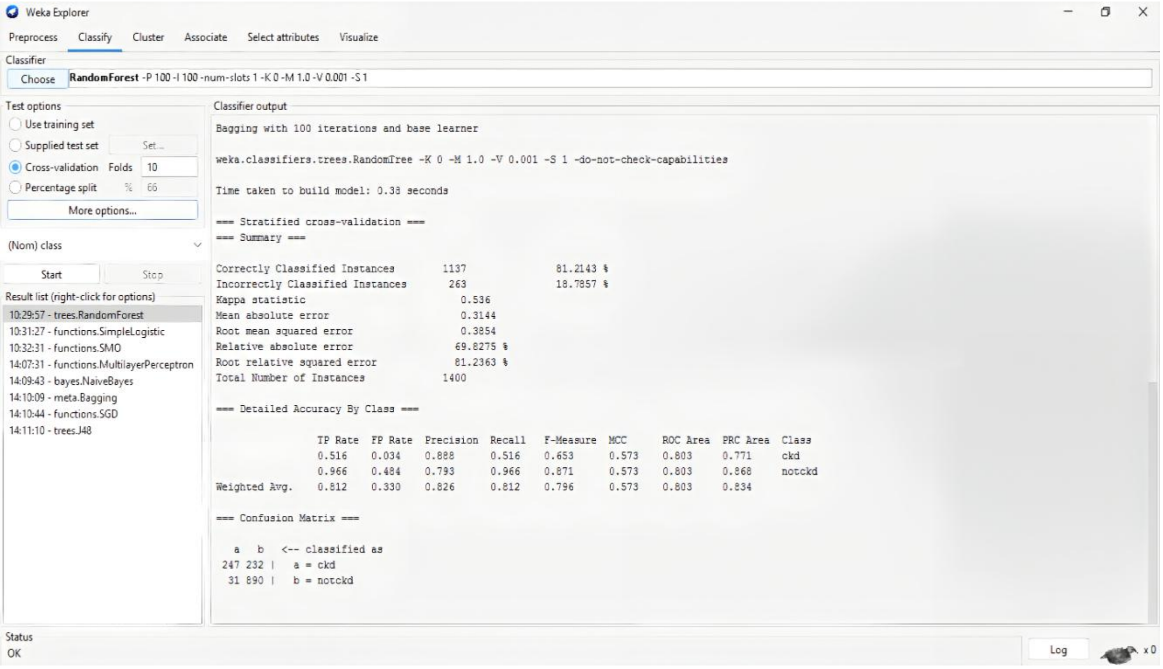


Figure 6: PCA Order of Importance

In the final experiments, the data mining algorithms are tested after applying PCA in order to acquire optimal accuracy with reduced features. The classifier Random Forest achieved 81.2857%, while Logistic Regression achieved 77.5741% and Support Vector Machine achieved 76.5741%.
Figure 7. below shows the result of Random Forest classifier, the model has an accuracy of 81.2967% with correctly classified instances of 1138, 18.7143% incorrectly classified instances of 262. The model took 0.38 seconds to build and an estimated time of 4.12 seconds to execute, mean absolute error of 0.3144, kappa statistics of 0.536, root mean squared error of 0.3844, relative absolute error of 69.82% and root relative squared Error of 81.23%. The figure also shows the confusion matrix of the model.

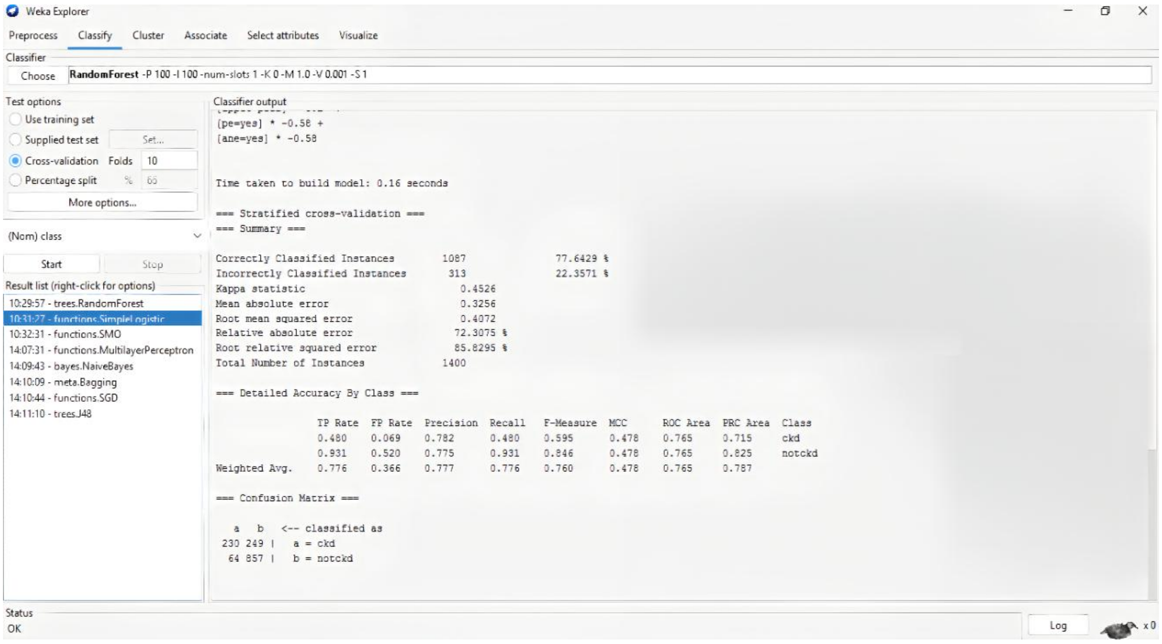


```
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
```

```
Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.38 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1137               81.2143 %
Incorrectly Classified Instances       263               18.7857 %
Kappa statistic                          0.536
Mean absolute error                      0.3144
Root mean squared error                  0.3854
Relative absolute error                 69.8275 %
Root relative squared error             81.2363 %
Total Number of Instances             1400

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.516    0.034    0.888      0.516    0.653      0.573    0.803     0.771     ckd
                 0.966    0.484    0.793      0.966    0.871      0.573    0.803     0.868     notckd
Weighted Avg.    0.812    0.330    0.826      0.812    0.796      0.573    0.803     0.834

=== Confusion Matrix ===

   a   b   <-- classified as
 247 232 |   a = ckd
  31 890 |   b = notckd
```

Figure 7: Random Forest Model Result

Figure 8. shows the Logistic Regression model's final results after the experiment; the model obtained an accuracy of 77.5714% with 1086 correctly classified instances at the end of the analysis, 22.4286% incorrectly classified instances of 314, 0.19 seconds to build the model and an approximate of 1.89 seconds execution time, mean absolute error of 0.3262, kappa statistics of 0.4518, root mean squared error of 0.408, relative absolute error of 72.4411%, and root relative squared error of 85.9923%. The figure also shows the confusion matrix of the model.

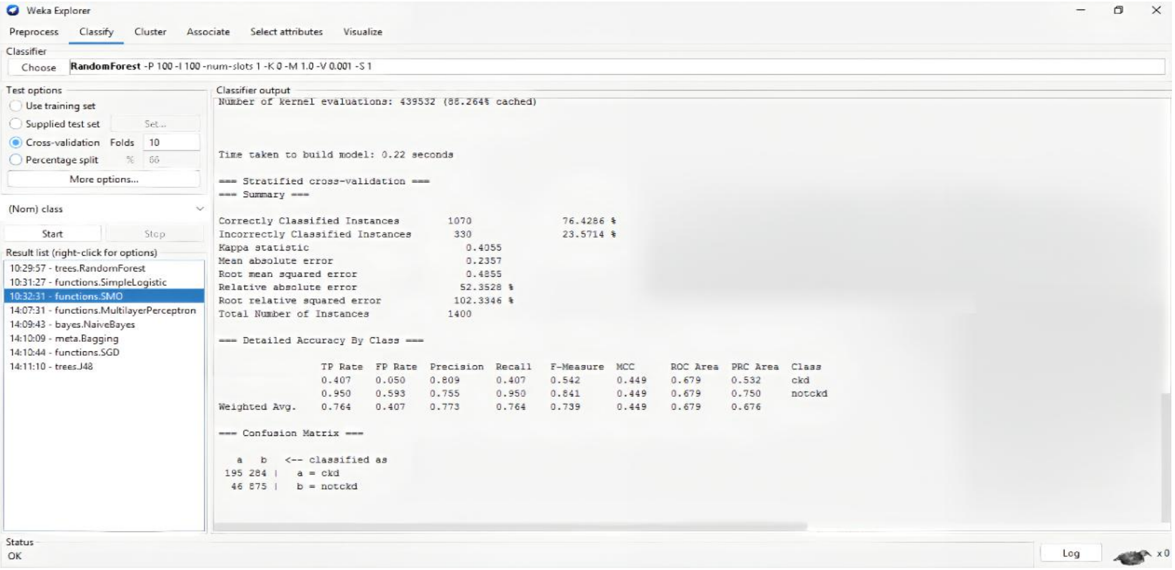


```
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
```

```
[pe=yes] * -0.58 +
[ane=yes] * -0.58


Time taken to build model: 0.16 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1087               77.6429 %
Incorrectly Classified Instances       313               22.3571 %
Kappa statistic                          0.4526
Mean absolute error                      0.3256
Root mean squared error                  0.4072
Relative absolute error                 72.3075 %
Root relative squared error             85.8295 %
Total Number of Instances             1400

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.480    0.069    0.782      0.480    0.595      0.478    0.765     0.715     ckd
                 0.931    0.520    0.775      0.931    0.846      0.478    0.765     0.825     notckd
Weighted Avg.    0.776    0.366    0.777      0.776    0.760      0.478    0.765     0.787

=== Confusion Matrix ===

   a   b   <-- classified as
 230 249 |   a = ckd
  64 857 |   b = notckd
```

Figure 8: Logistic Regression Result

Figure 9. shows the support vector machine model's final results after the experiment; the model achieved an accuracy of 76.5714% with 1072 correctly classified instances at the end of the analysis, 23.4286% incorrectly classified instances off 328, the model took 0.22 seconds to build and an approximate of 2.11second execution time, mean absolute error of 0.2343, kappa statistics of 0.4078, root mean squared error of 0.484, relative absolute error of 52.0355%, and root relative squared error of 102.024%. The figure also shows the confusion matrix of the model.



```
Classifier
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options
○ Use training set
○ Supplied test set
● Cross-validation  Folds 10
○ Percentage split  % 66

(Nom) class

Result list (right-click for options)
10:29:57 - trees.RandomForest
10:31:27 - functions.SimpleLogistic
10:32:31 - functions.SMO
14:07:31 - functions.MultilayerPerceptron
14:09:43 - bayes.NaiveBayes
14:10:09 - meta.Bagging
14:10:44 - functions.SGD
14:11:10 - trees.J48

Status
OK
```

```
Classifier output
Number of kernel evaluations: 439532 (86.264% cached)



Time taken to build model: 0.22 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1070               76.4286 %
Incorrectly Classified Instances       330               23.5714 %
Kappa statistic                          0.4055
Mean absolute error                      0.2357
Root mean squared error                  0.4855
Relative absolute error                 52.3528 %
Root relative squared error            102.3346 %
Total Number of Instances             1400

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.407    0.050    0.809      0.407    0.542      0.449    0.679     0.532     ckd
                 0.950    0.593    0.755      0.950    0.841      0.449    0.679     0.750     notckd
Weighted Avg.    0.764    0.407    0.773      0.764    0.739      0.449    0.679     0.676

=== Confusion Matrix ===

   a   b   <-- classified as
 195 284 |   a = ckd
  46 875 |   b = notckd
```

Figure 9: Support Vector Machine

Table 3 summarizes clearly the result of the three selected algorithms that were used to predict the early stage of CKD, using 15 attributes and number instances of 1400. it shows that Random Forest has predicted correctly higher number of classified instances of 1137, with a low rate of incorrectly predicted instances of 263, Logistic regression marks the fastest in building the model with approximately (0.16sec). SVM is last in time building the model with approximately (0.22sec). The number of correctly classified instances indicates the accuracy of the model; therefore Table 3 concludes that RF performs better in accuracy than LR and SVM, but LR is fastest in building the model, followed by SVM then RF being the slowest in time building the model.

Table 3: Performance of the Classifiers

Evaluation Criteria	Classifiers		
	RF	**LR**	**SVM**
Time to Build Model (in sec)	0.38	0.16	0.22
Correctly Classified	1137	1087	1070
Incorrectly Classified	263	313	330
Precision	0.892	0.778	0.777
	0.793	0.775	0.755
Recall	0.516	0.482	0.776
	0.967	0.928	0.953
Result (%)	81.2967	77.5714	76.5714

Experimental Results of 10-Fold Cross Validation

In the experiment, it is observed that in Table 4. Random Forest model has the highest accuracy value with (0.8128), (0.5156) sensitivity and (0.9674) specificity. Which shows that RF is the best performing model in predicting the early stage of chronic kidney disease data, SVM has the lowest accuracy value and performs less and slower.

Table 4: Experimental Result of 10-Fold Cross Validation of all Models and Confusion Matrix

Classification Technique	Confusion Matrix		Accuracy	Sensitivity	Specificity
RF	A	B	0.8128	0.5156	0.9674
	247	232			
	30	891			
LR	231	248	0.7757	0.4822	0.9283
	66	855			
SVM	194	285	0.7657	0.4050	0.9533
	43	**878**			

From the above Table 4. the confusion matrix of RF, true positive for class a=’ckd’ is 247, while false positive is 232, whereas for class b= ’notckd’ is 30 and false positive 891. The diagonal element of correct matrix is 247+891 = 1138 and other instances 232+30 = 262 represent the incorrect instances.

Performance Evaluation of Rf and The Other Algorithms

The models were similarly built using the preprocessed 10-fold cross validation dataset. The models are also trained and tested using 10-fold Cross Validation and were evaluated with other performance evaluation metrics. The performance metrics result of each trained model; RF, LR and SVM have been presented with feature selection method. The Models were first trained and tested with all features and then we apply the PCB feature selection method. The 10-fold cross validation performance metrics results for three classifiers of multiclass dataset. The accuracy is 81% from RF was obtained and is said to have a larger number of instances and also applying feature selection of PCB to have a reduced number of attributes from 25 to 15. Other classification models were tested, but achieved lower accuracy as seen in Table 6. it shows the level of accuracy of other models compared to that of RF, which tell the RF model is best in predicting the early stage of CKD using this dataset.

Table 5: Evaluation Between RF and Other Models

SN	Classifier	Dataset	Method	Result (%)
1	MLP	CKD DATA	FEATURE SELECTION	75
2	NAÏVE BAYES	CKD DATA	FEATURE SELECTION	77.371
2	BAGGING	CKD DATA	FEATURE SELECTION	77.9286
3	J48	CKD DATA	FEATURE SELECTION	74.1427
4	RF	CKD DATA	FEATURE SELECTION	81.2967
5	AdaBoost	CKD DATA	FEATURE SELECTION	74.1429

The table 5. shows that from the experiment, RF performs best in accuracy for predicting the early stages of ckd compared to other models while using the same ckd dataset.

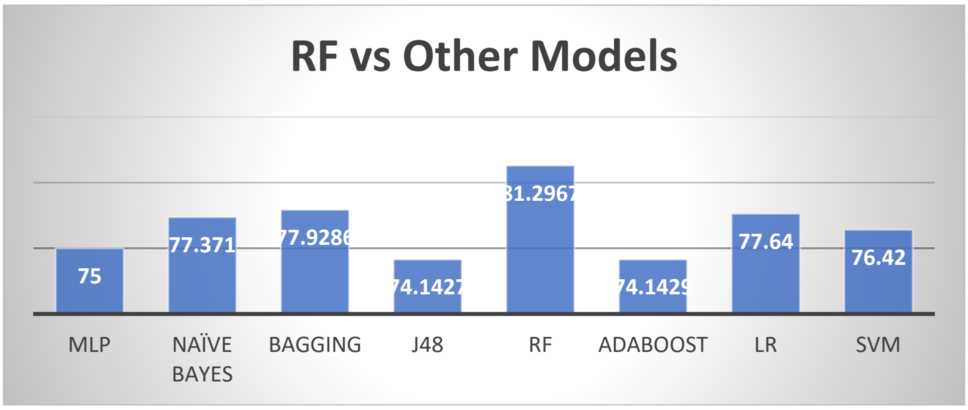


Figure 10: Visual Representation of RF with other Models

CONCLUSION

The findings provides valuable insight for identify patients at risk of developing CKD at an early stage where early detection allows for timely intervention and treatment, potentially slowing or halting the progression of the disease by employing reduced test features from an Electronic Health Dataset using Machine learning (ML) techniques, which allows healthcare providers to implement preventive measures and interventions promptly. By predicting CKD development, healthcare professionals can strive to improve patient outcomes, reduce complications, and enhance overall quality of life. This research can potentially identify new avenues for interventions and guide the development of novel therapeutic strategies.

REFERENCES

B. Khan, R. Naseem, M. Ali, M. Arshad, and N. Jan, ‘‘Machine learning approaches for liver disease diagnosing,’’ *Int. J. Data Sci. Adv. Anal.*, vol. 1, no. 1, pp. 27–31, 2019

Caragea, D. Cook, H. Wickham, and V. Honavar, “Visual methods for examining svm classifiers,” in *Visual Data Mining*.

Chen, T. K., Knicely, D. H. & Grams, M. E. (2019). Chronic kidney disease diagnosis and management. *JAMA* 322, 1294.

Gudeti, B., Mishra, S., Malik, S., Fernandez, T. F., Tyagi, A. K., & Kumari, S. (2021). A noval approach to predict chronic kidney disease using machine learning algorithms.

Iliyas I. I., Saidu I. R., Ali B. D., & Suleiman T. (2020). Prediction of chronic kidney disease using deep neural network. *FUDMA Journal of Sciences (FJS)* Vol. 4 No. 4, December, 2020, pp 34 – 41 DOI: https://doi.org/10.33003/fjs-2020-0404-309

Ilyas, H., Ali, S., Ponum, M.1., Osman Hasan, O., Tahir M. M., Iftikhar, M., & Hussain M. M (2021) Chronic kidney disease diagnosis using decision tree algorithms. BMC Nephrology (2021) 22:273 https://doi.org/10.1186/s12882-021-02474-z

Islam, A., Majumder, Z.H., Hussein, A. (2023). Chronic kidney disease prediction based on machine learning algorithms. *Journal of Pathology Informatics 14 100189*

Kempf-Leonard, K. (2004) “Encyclopedia of social measurement,”.

Khan, B., Naseem, R., Muhammad F., Abbas, G, & Kim, S. (2020) An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Prophecy. *IEEE Access* https://doi.org/10.1109/ACCESS.2020.2981689 V8

MacGregor, J. (2013). Predictive Analysis with SAP®. Bonn: Galileo Press.

National Kidney Foundation Inc. (2024) How kidneys work. Retrieved from https://www.kidney.org/kidneydisease/howkidneyswrk

Rahman, M., Islam L, Rana, M., Zannat T. M., Ferdous, S. J., & Tasmia, A. S. (2022) A Predictive Analysis of Chronic Kidney Disease by Exploring Important Features. *International Journal of Computing and Digital Systems 11, No.1, 167-176 (Jan-2022).*

Sharma, N., Singh, A., (2019). Diabetis detection and prediction using Machine Learning/IoT: a survey: ICAICR . CCIS, vol. 955, pp. 471-479. Springer Singapore. https://doi,0rgs/10.1007/978-981-130314004_42 Springer, 2008, pp. 136–153.

Singh, V., Asari, V.K., Rajasekaran, R. (2022) A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. *Diagnostics 2022, 12, 116*. https://doi.org/10.3390/diagnostics12010116

Srivastava, N., Lamba, T., Agwarwal, M. (2020). Comparative analysis of different Machine Learning Techniques. Futuristic Trends in Networks and Computing Technologies (pp.245-255)

Zahid, U., & Mona, J., (2023) Early Detection and Diagnosis of Chronic Kidney Disease Based on Selected Predominant Features. Hindawi Journal of Healthcare Engineering Volume 2023, Article ID 3553216, 8 pages https://doi.org/10.1155/2023/3553216