Academy Journal of Science and Engineering
Available online @ www.academyjsekad.edu.ng
AJSE 15(1) 2021
Original Research

Comparative Analysis of Some Prominent Machine Learning Algorithm for the Prediction of Chronic Kidney Disease

[1]Iliyas, I.I., [2] Isah, R.S., [3]Ali, B.D. and [4]Andra, U.

[1,3,4]Department of Mathematical Science, University of Maiduguri
[2]Department of Cyber Security, Nigerian Defence Academy, Nigeria

Abstract

Chronic Kidney Disease (CKD) is a disorder against proper function regarding kidneys, as kidneys filter our blood whenever CKD gets worse, our blood receives wastes at a higher level, which results in sickness. It also has a substantial financial problem for families of subjects with a medical issue in Nigeria. Among the necessary measures that need action concerning the increase of CKD is detecting the disease early and with different data mining techniques.
Data mining is gradually becoming more prevalent nowadays in healthcare, as also in fraud, abuse detection etc. Classification is a more useful data mining function to handle items in a collection to class or target categories. For obtaining essential information from medical database, machine learning and statistical analysis can assist in extracting hidden patterns and identify relationships from vast among of data. In this study, we compared five (5) different models namely: Deep Neural Network (DNN), Artificial Neural Network (ANN), Naïve Bayes (NB), Logistic Regression (LR), and K-Neighbor Nearest (KNN) to predict CKD on Gashua General Hospital (GGH) dataset. The study achieved an accuracy of 98% for DNN, KNN: 96%, NB: 97%, LR: 96% and ANN: 96%. The best performance was obtained with DNN with the highest accuracy and can be applied in real world application.

Key Words: Chronic Kidney Disease, Machine Learning, Deep Neural Network, Artificial Neural Network, Naives Bayes, Logistics Regression, K-Neighbor Nearest

Corresponding author's e-mail: iliyasibrahimiliyas@uniamaid.edu.ng

1. INTRODUCTION

Chronic Kidney Disease (CKD) is a disorder against proper function regarding kidneys. Our kidneys balance the salt and minerals such as calcium, phosphorus, sodium, and potassium that circulate in our blood, filter wastes from the blood, and remove them through urination. This filtering process includes excess fluids from our body (Chukwuonye et al., 2018). As kidneys filter our blood whenever CKD gets worse, our blood receives wastes at a higher level, which results in sickness. Reduced Glomerular Filtration Rate (GFR) increased urinary albumin excretion, both GFR and urinary albumin are the binding definition terms for Chronic Kidney Disease (CKD). CKD was ranked from the list of diseases that cause global deaths in the 1990s; by 2010, it had fallen to 28th in the list of global death (Chukwuonye et al., 2018). Levey et al. (2007) made it clear that the level by which it rises is noted to be second only to HIV & AIDs. According to Luyckx & Stanifer, (2018), CKD increased globally from 19 million in 1990 to 33 million in 2013, and in 2010, 2.62 million individuals got dialysis around the world. The requirement for dialysis predicted to twofold by 2030. With the attention being paid globally to CKD is inferable to five variables: the quick increment in its predominance, the gigantic fetched of treatment, later information demonstrating that direct illness is the tip of an ice sheet of undercover infection, an appreciation of its significant part in expanding the chance of cardiovascular disease, and the revelation of successful procedures to anticipate its movement (Barsoum, 2006).

In Nigeria, the situation is such that CKD represents about 8-10% of hospital admissions (Ulasi & Ijoma 2010). An investigation was carried out at the University of Maiduguri Teaching Hospital (UMTH) and found that approximately 15% of the individuals who come to the clinic from the catchment zones have kidney sickness, and 20 out of 100 patients are from Bade community (Gashua) of Yobe State (Ummate et al, 2008). It also has a substantial financial problem for families of subjects with a medical issue in Nigeria. Among the necessary measures that need action concerning the increase of CKD is detecting the disease early and with different data mining techniques. Data mining is gradually becoming more prevalent nowadays in healthcare, as also in fraud, abuse detection etc. (Iliyas et al, 2021). Classification is a more useful data mining function to handle items in a collection to class or target categories. By obtain essential information from medical database, machine learning techniques, statical analysis and dataset has shown tremendous success in extracting hidden patterns and identify relationships from vast among of data (Padmanaban & Parthiban 2016). Exploring many machine learning models for the prediction of kidney disease is important because various models have their own way of identifying patterns on dataset, but by comparing more than one model, it helps in knowing which model can predict well than the other.

In this study, we compared five (5) different models namely: Deep Neural Network (DNN), Artificial Neural Network (ANN), Naïve Bayes (NB), Logistic Regression (LR), and K-Neighbor Nearest (KNN) to predict CKD on Gashua General Hospital (GGH) dataset. Performance evaluation of the model was computed by computing the accuracy, Recall, Precision, and F1 Score.

Literature Review

Norouzi et al. (2016) Presented an Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict the renal failure time frame of CKD on a 10-year real clinical data of diagnosed patients. The dataset had 10 attributes, in their preprocessing steps, they replaced missing values with the mean value of the column with a missing value. ANFIS model was used for estimating GFR at subsequent 6, 12, or 18 months. Their model was able to achieve an accuracy of 95%. features/attribute but did not validate ANFIS models on a reduced feature for prediction.

Sathya & Suresh (2018) Employed DT and NB as a machine learning algorithm to predict CKD using UCI machine repository dataset with 25 attributes and achieved an accuracy of 99.25% and 98.75% for DT and NB respectively, showing DT as a better algorithm in terms of predicting the presence and absence of CKD.

Chimwayi et al. (2017) applied the use of a neuro-fuzzy algorithm as a technique to predict the risk of CKD patients, using a UCI dataset which had 25 attributes/features (11 numeric and 14 nominal), with an accuracy of 100%, sensitivity of 100% and specificity of 97%, they suggested that their work can be added in the domain of healthcare and also can be used in making it easier for professionals in diagnosing, treating patients and identifying relations of diseases suffered by patients.

Arasu & Thirumalaiselvi (2017) used Weighted Average Ensemble Learning Imputation (WAELI) to perform feature selection and predicted CKD with the selected features on the UCI dataset, the algorithm used by the authors in the prediction is: SVM and ANN with an accuracy of 73% for both algorithms while after feature selection, an accuracy of 78% for both ANN and SVM was achieved

Misir et al. (2017) Predicted CKD and NCKD with reasonable accuracy using a lesser number of features on a balanced dataset gotten from the UCI repository dataset, they performed feature extraction and reduction using CFS, with WEKA as a tool, their work was able achieved promising accuracy with the use of two classifiers namely: Correlation-based feature subset selection and Levenberg–Marquardt on 8 attributes.

Arafat et al. (2018) Studied an automated detection of CKD with clinical data using RF and NB based on a comparative study on the UCI dataset, they computed the weight of each attribute used in the dataset. Their result shows that RF has higher accuracy of 98%, followed by LR and NB with 96% for each.

Alshebly & Ahmed (2019) applied different machine learning algorithm, which are ANN and LR, to a problem in the domain of medical diagnosis and analysed their efficiency of the prediction on the University of California Irvine (UCI) dataset with 153 cases and 11 attributes of CKD patients, the observed performance of the ANNs classifier is better than LR mode with the accuracy of 84.44%, sensitivity of 84.21, specificity of 84.61% and Area Under the Curve (AUC) of 84.41% and found that the most critical factors that have a clear impact on CKD patients are creatinine and urea, they ignored cases with missing values and only used 153 cases.

Ayon & Islam (2019) Proposed a strategy for the diagnosis of Diabetes using DNN on the PIM Indian Diabetes (PID) dataset from UCI, they replaced missing values with the column mean, they achieved an accuracy of 98.35%, F1 Score: 98% and MCC: 97% for five-fold cross-validation. Additionally, 97.11% accuracy, Sensitivity: 96.25% and Specificity: 98.80% obtained for ten-fold cross-validation and indicated that five-fold cross-validation showed better performance.

Kriplani et al. (2019) Used 224 records of CKD that were gotten from a dataset online called UCI machine learning repository namely; chronic kidney diseases going back to the year 2015, and proposed an algorithm, they did not explain preprocessing steps taken in their work. Their method is based on a deep artificial neural network, which predicts whether a patient has CKD or NCKD with 97% accuracy. Compared to other available algorithms, their model shows better results, which was validated using the cross-validation technique

Sharma & Parmar (2020) Proposed a model for heart disease prediction with a DNN model on heart disease UCI dataset with six (6) different classifiers KNN, SVM, NB, RF and DNN using Talos optimization. Their work indicated an accuracy for KNN:90.16%, LR: 82.5%, SVM: 81.97%, NB: 85.25% and DNN with Talos optimization: 90.78%.

Various work had shown tremendous result but there exist few considerations of implementing deep learning models for prediction, that is the reason for the comparative analysis using both traditional machine learning and deep learning algorithm, the study tried to use a larger number of dataset compare to other literature to determine if a good accuracy can be achieved with larger number of instances, and finally the study evaluated the performance of the models based on accuracy, precision, F1 score and recall score.

Methodology

Data Collection

The main test conducted to determine chronic kidney disease is either through blood test, urine test or image scans. This study focused on designing a model to predict CKD's presence and absence in humans from the age of 5 years to 90 years old in GGH Yobe State, Nigeria, the datasets is from the year 2018 to 2019. The data was not in a digital form. The record was inputted into Microsoft Excel and saved as CSV format. The dataset contains some missing values but were replaced with the mean values of the cells. The dataset contains 1200 patients records with 600 samples for CKD and 600 for NCKD cases. It had 11 attributes/features: Age, Gender, Sodium, Potassium, Chloride, Bicarbonate, Urea, Creatinine, Uric Acid, Albumin, and Classification including a target variable classified into a binary classification of CKD and NCKD as shown in Table 1. The data classifier to be used are DNN, KNN, LR, NB, and ANN which will employ supervised learning.

Table 1: Dataset Attributes

No.	Attribute Name	Coding of Attribute	Types of Attribute
1.	Sex (Gender)	1 Male 0 Female	Nominal
2.	Age (Age)	NA	Numeric
3.	Sodium (Sod)	NA	Numeric
4.	Potassium (Pot)	NA	Numeric
5.	Chloride (Chl)	NA	Numeric
6.	Bicarbonate (Bica)	NA	Numeric
7.	Urea (Urea)	NA	Numeric
8.	Uric Acid (UA)	NA	Numeric
9.	Albumin (Alb)	NA	Numeric
10.	Creatinine (Crea)	NA	Numeric
11.	Classification	1 KD 0 NKD	Nominal

Gender: Gender is a factor in developing CKD as men have a high tendency of having CKD than females.

Age: Age is a factor in developing CKD as the decaying of kidney function accelerates as people get older. It is of numerical value in the data set.

Albumin: Albumin is a substance that is often found in the urine if the kidney has disruptive functionality. Albumin is a protein-based substance which should not be present in the urine of a healthy person, albumin level from 30 and above could point to kidney problems. In this, the albumin attribute is numeric.

Urea: This is the Urea Nitrogen level in the blood. A healthy kidney Separates and discharges the urea nitrogen through urine. A high level of blood urea means the kidney is filtering the urea nitrogen properly. This dataset has a numerical value.

Creatinine: Creatinine clearance in urine is measured to estimate the GFR rate of the kidney. Serum creatinine is measured in mmol/L and it is numerical in the dataset.

Sodium: A high-level salt diet can alter sodium balance, triggering the kidneys to reduce functioning and removes less water resulting in higher blood pressure, it is measured in gram(g). It is of numerical value.

Potassium: When the kidney failed, it can no longer remove access to potassium, so the level of potassium build-ups in the body. High potassium can cause advanced stages of CKD, it is measured in millimoles per liter (mm/l). This dataset has a numerical value.

Chloride: Chloride is used to remove acid from the blood, high chloride shows a sign of CKD, chloride level above 106 could trigger kidney problems. This dataset has a numerical value.

Bicarbonate: it preserves renal function in excremental chronic kidney disease, a level between 24 and 26 mEq/L could point to kidney problem, low bicarbonates shows a sign of CKD. It is a numeric attribute.

Uric Acid: is build up from urate crystal, high Uric acid causes CKD. This dataset has a numerical value.

Classification: classification attribute is used to classify either "CKD" referring to having chronic kidney disease and "NOT CKD".

Machine Learning Models

Deep Neural Network (DNN)

A Deep Neural Network (DNN) is a form of deep learning technique that comprises an input layer, several hidden layers, and an output layer. Each layer comprises several units called neurons. These neurons are also referred to as artificial neurons. A neuron obtains several inputs, performs a weighted summation over its inputs with a bias, then the resulting sum goes activation process with an activation function to yield output. Each neuron contains a vector of weights associated with its input size and a bias that should be optimized during the training process (Chahal & Gulia, 2019).

Naive Bayes (NB)

Naive Bayes is a machine learning technique or classifier which is based on Bayes theorem that has independent assumption between features. The one-dimensional Naive Bayes classifier computes the ratio of the log probabilities of the features belonging to all the classes. The naive Bayes classifier computes the class value probability assuming each of the attributes independently. This means Naive Bayes does not consider correlation that is in-between attributes. Naive Bayes is a very scalable classifier, but it can create a bias towards one or more attributes which often results in inaccuracy (Arafat et al, 2018). Equation (1) shows NB formula.

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Logistic Regression (LR)

A logistic regression has a gaussian distribution which possess odd ratio, were a log odd of the input variable (disease status) is modelled as linear combination of target variables. LR is suitable for binary target variables, the disease status of breast cancer patient between Benign and Malignant is replaced with 1/0 in the breast cancer dataset during model evaluation with LR (Ganggayah et al., 2019). LR is represented in equation (2)

$$p = \frac{e^{a+bx}}{1 + e^{a+bx}} \quad \ldots (2)$$

K-Nearest Neighbor (K-NN)

K-Nearest Neighbour algorithm is the supervised learning algorithm is used for statistical estimation and pattern recognition. K-NN is a lazy learner because it differs from the classifiers previously. The entire memory will store in and at a time it will take out and move to a based class. The presented training data is simply stored when a new query instance is encountered, the related instances are retrieved from memory and used to classify the new query instance. Hamming distance is used for categorical data while Euclidean distance is used for continuous data. All the training samples are calculated using any distance measure which is Euclidean and Hamming (Arafat et al., 2018).

Artificial Neural Networks

ANN is a numerical show that tries to recreate the structure and functionalities of natural neural systems. The basic building block of each artificial neural network is the ANN, a simple mathematical model (function). This model must have three simple sets of rules: summation, multiplication, and activation. In the artificial neuron entrance, the inputs are weighted, meaning that every input value is multiplied with individual weight. The middle section of an artificial neuron is a sum function that sums all weighted inputs and bias. At the exit of an artificial neuron, the sum of previously weighted inputs and bias passes through an activation function called the transfer function (Krenker et al., 2011).

When there exist two or more artificial neurons, we get an artificial neural network. When a single artificial neuron has no use in solving real-life problems, artificial neural networks have it. ANN can be used to solve composite real-life problems by using their basic blocks (artificial neurons) to process information in a distributed, non-linear, parallel, and local way.

Performance Metrics

Confusion Matrix: confusion matrix indicates the model's statistical suitability and its compatibility with the dataset. It can be defined as a table layout that is used explicitly for the visualization of algorithm performance (Alshebly & Ahmed, 2019); Table 2 provided a summary of confusion matrix.

Table 2: Confusion Matrix

Classification		Observation	
		Negative	Positive
Positive	Negative	True Negative (TN)	False Positive (FP)
	Positive	False Negative (FN)	True Positive (TP)

i. **Accuracy**- It is used to classify the number of correctly predicted data points out of all data points. it is the number of data points that were predicted correctly divided by the total number of data points prediction made (Iliyas et al, 2021). It is expressed in equation (3):

$$\boldsymbol{Accuracy = \frac{TN + TP}{TP + FP + FN + TN} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)}$$

ii. **Precision**: It is defined as the portion of relevant instances among the retrieved instances. It is given as the correlation number between the correctly classified modules to entire classified fault-prone modules (Alshebly & Ahmed, 2019). It is expressed in equation 4:

$$\boldsymbol{Precsion = \frac{TP}{TP + FP} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)}$$

iii. **Recall/ Sensitivity**: Recall is a metric that measures the number of correct positive classified data points made out of all the positive data points that are supposed to be made (Alshebly

& Ahmed, 2019). It is expressed in equation 5:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP + FN}} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.\dots..(\boldsymbol{5})$$

iv. **F1 Score**: This use to determine the mean between precision and recall. It describes the preciseness (how many records can be correctly classified by the model) and robustness (it avoids missing any significant number of records) of a model (Alshebly & Ahmed, 2019). The expression of f1-score is in equation 6:

$$\boldsymbol{F1\ Score} = \boldsymbol{2} \times \frac{\boldsymbol{1}}{\frac{\boldsymbol{1}}{\boldsymbol{precision}} + \frac{\boldsymbol{1}}{\boldsymbol{recall}}} \dots\dots\dots\dots\dots\dots\dots\dots...(\boldsymbol{6})$$

Results and Discussion

From the analysis of different prediction models, it has been observed that DNN model proved to be more reliable in the prediction of CKD, this section provided the summary of the results that was achieved by the five models. Figure 1 depicts the precision, recall, f1 score and accuracy of DNN and Table 3 depicts the confusion matrix of DNN, Figure 2 depicts the precision, recall, f1 score and accuracy of KNN and Table 4 depicts the confusion matrix of KNN, Figure 3 depicts the precision, recall, f1 score and accuracy of NB and Table 5 depicts the confusion matrix of NB, Figure 3 depicts the precision, recall, f1 score and accuracy of LR and Table 6 depicts the confusion matrix of LR, Figure 4 depicts the precision, recall, f1 score and accuracy of ANN and Table 7 depicts the confusion matrix of ANN, and Table 8 summarized the comparison results of all the models.

```
Classification report -
               precision    recall  f1-score   support

           0       0.99      0.98      0.98        94
           1       0.98      0.99      0.98        86

    accuracy                           0.98       180
   macro avg       0.98      0.98      0.98       180
weighted avg       0.98      0.98      0.98       180
```

Figure 1: Deep Neural Network Results

During experiments shown in Figure 1, data used for training was 70%, while the dataset from testing was 30% which amounted to an accuracy of 98% with a precision of 0.99 for NCKD and 0.98 for CKD, recall of 0.98 for NCKD, and 0.99 for CKD, f1 score of 0.98 for NCKD and 0.98 for CKD, support of 94 for NKD and 86 for NCKD.

Table 3: Deep Neural Network Confusion Matrix

		Negative	Positive
	Negative	93	1
	Positive	0	86

As shown in Table 3, DNN produced a good confusion matrix which shows that from the 30% CKD dataset used for testing, ninety-three (93) were true negative i.e they were predicted to be correctly NCKD and one (1) of them were false negatives, meaning they were wrongly predicted to be NCKD while they are CKD, and also indicates that from 30% of NCKD dataset used for testing, zero (0) of them was false positive, meaning it was predicted to be wrongly CKD while eighty-six (86) of the dataset were true positive, meaning they were correctly predicted NCKD.

```
[[84  3]
 [ 4 89]]
              precision    recall  f1-score   support

           0       0.95      0.97      0.96        87
           1       0.97      0.96      0.96        93

    accuracy                           0.96       180
   macro avg       0.96      0.96      0.96       180
weighted avg       0.96      0.96      0.96       180
```

Figure 2: K-Neighbor Nearest Results

During the experiment shown in Figure 2, KNN model, data used for training was 70%, while dataset for testing was 30%, which amounted to an accuracy of 96%, with a precision of 0.95 for NCKD and 0.97 for CKD, recall of 0.97 for NCKD and 0.96 for CKD, f1 score of 0.96 for NCKD and 0.96 for NCKD, support of 87 for NCKD and 93 for CKD.Table 4: KNN Confusion Matrix

Table 4: K-Neighbor Nearest Confusion Matrix

		Negative	Positive
	Negative	**84**	**3**
	Positive	**4**	**89**

As shown in Table 4, the KNN model produced a good confusion matrix which shows that from the 30% of the CKD dataset used for testing, eighty-four (84) were true negative i.e they were predicted to be

correctly NCKD and three (3) of which were false negative, meaning they were wrongly predicted to be NCKD while they were NCKD, and also indicated that 30% of the CKD dataset used for testing, four (4) of them were false positive meaning they were predicted to be wrongly CKD while eighty-nine (89) of the dataset were true positive meaning they were correctly predicted to be CKD.

```
Classification report -
              precision    recall  f1-score   support

           0       0.99      0.95      0.97        88
           1       0.96      0.99      0.97        92

    accuracy                           0.97       180
   macro avg       0.97      0.97      0.97       180
weighted avg       0.97      0.97      0.97       180
```

Figure 3: Naïve Bayes Results

During the experiment shown in Figure 3, NB model, data used for training was 70%, while dataset for testing was 30%, which amounted to an accuracy of 97%, with a precision of 0.99 for NCKD and 0.96 for CKD, recall of 0.95 for NCKD and 0.99 for CKD, f1 score of 0.97 for NCKD and 0.97 for NCKD, support of 88 for NCKD and 92 for CKD.

Table 5: Naïve Bayes Confusion Matrix

		Negative	Positive
	Negative	84	1
	Positive	4	91

As shown in Table 5, the NB model produced a good confusion matrix which shows that from the 30% of the CKD dataset used for testing, eighty-four (84) were true negative i.e they were predicted to be correctly NCKD and one (1) of which were false negative, meaning they were wrongly predicted to be NCKD while they were NCKD, and also indicated that 30% of the CKD dataset used for testing, four (4) of them were false positive meaning they were predicted to be wrongly CKD while ninety-one (91) of the dataset were true positive meaning they were correctly predicted to be CKD.

```
Classification report -
              precision    recall  f1-score   support

           0       0.92      1.00      0.96        90
           1       1.00      0.91      0.95        90

    accuracy                           0.96       180
   macro avg       0.96      0.96      0.96       180
weighted avg       0.96      0.96      0.96       180
```

Figure 4: Logistic Regression Results

During the experiment in Figure 4, LR model, data used for training was 70%, while dataset for testing was 30%, which amounted to an accuracy of 96%, with a precision of 0.92 for NKD and 1.00 for CKD, recall of 1.00 for NCKD and 0.91 for CKD, f1 score of 0.96 for NCKD and 0.95 for NCKD, support of 90 for NCKD and 90 for CKD.

Table 6: Logistic Regression Confusion Matrix

		Negative	Positive
	Negative	84	6
	Positive	4	86

As shown in Table 6, the LR model produced a good confusion matrix which shows that from the 30% of the CKD dataset used for testing, eighty-four (84) were true negative i.e they were predicted to be correctly NCKD, and six (6) of which were false negative, meaning they were wrongly predicted to be NKD while they were NCKD, and also indicated that 30% of the CKD dataset used for testing, four (4) of them were false positive meaning they were predicted to be wrongly CKD while eighty-six (86) of the dataset were true positive meaning they were correctly predicted to be CKD.

.

```
Classification report -
              precision    recall  f1-score   support

           0       0.92      1.00      0.96        89
           1       1.00      0.91      0.95        91

    accuracy                           0.96       180
   macro avg       0.96      0.96      0.96       180
weighted avg       0.96      0.96      0.96       180
```

Figure 5: Artificial Neural Network Results

During the experiment in Figure 5, ANN model, data used for training was 70%, while dataset for testing was 30%, which amounted to an accuracy of 96%, with a precision of 0.92 for NCKD and 1.00 for CKD, recall of 1.00 for NCKD and 0.91 for CKD, f1 score of 0.96 for NCKD and 0.95 for NCKD, support of 89 for NCKD and 91 for CKD.

Table 7: ANN Confusion Matrix

		Negative	**Positive**
	Negative	**80**	**9**
	Positive	**1**	**90**

As shown in Table 7, the ANN model produced a good confusion matrix which shows that from the 30% of the CKD dataset used for testing, eighty (80) were true negative i.e they were predicted to be correctly NCKD and nine (9) of which were false negative, meaning they were wrongly predicted to be NCKD while they were NCKD, and also indicated that 30% of the CKD dataset used for testing, one (1) of them were false positive meaning they were predicted to be wrongly CKD while ninety (90) of the dataset were true positive meaning they were correctly predicted to be CKD.

Table 8: Results Comparison

Model	Class of Chronic Kidney Disease (CKD)	Accuracy	Precision	Recall	F1 Score	Support
Deep Neural Network (DNN)	NCKD	98%	0.99	0.98	0.98	94
	CKD		0.98	0.99	0.98	86
K Nearest Neighbors (KNN)	NCKD	96%	0.95	0.97	0.96	87
	CKD		0.97	0.96	0.96	93
Logistic Regression (LR)	NCKD	96%	0.92	1.00	0.96	90
	CKD		1.00	0.91	0.96	90
Naïve Bayes (NB)	NCKD	97%	0.99	0.95	0.97	88
	CKD		0.96	0.99	0.97	92
Artificial Neural Network (ANN)	NCKD	96%	0.92	1.00	0.96	89
	CKD		1.00	0.91	0.95	91

Table 8 highlighted the summary of the comparisons of the five (5) models namely: ANN, LR, NB, DNN and NB that was used for the prediction of CKD, including the accuracy results, the precision, recall and F1 Score of the predicted NCKD and CKD.

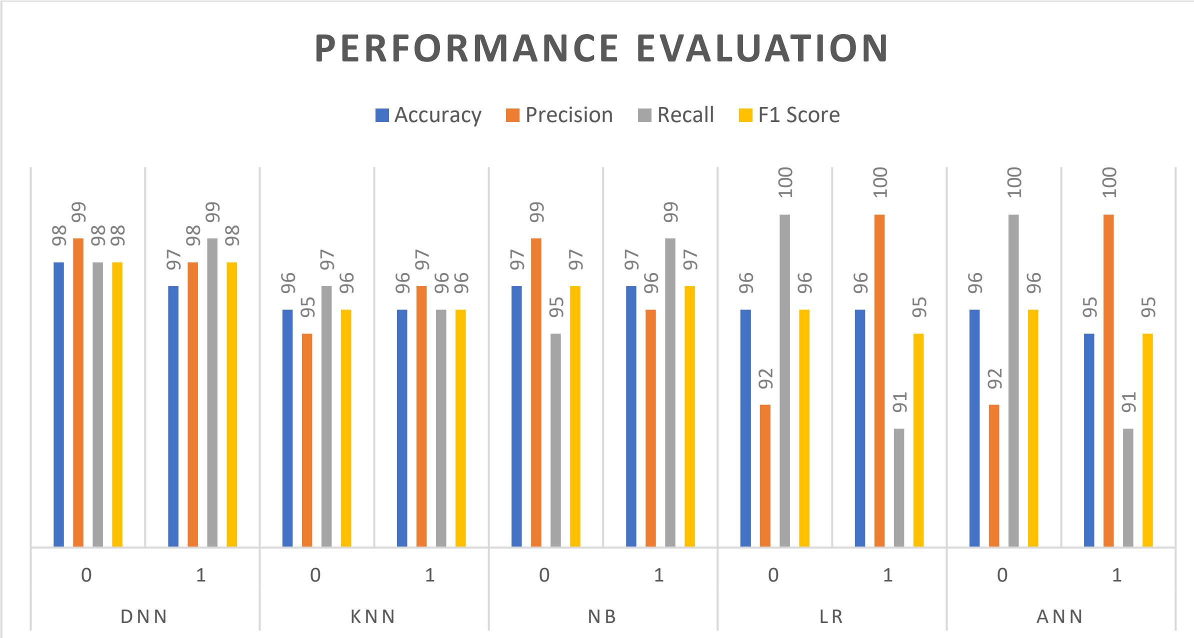

Figure 6: Chart Showing the Performance of the Models Result

From the results gotten in table 6 and chart displayed in Figure 6 which shows various performance levels, it indicates that DNN produced the best performance accuracy with an accuracy of 98%. Therefore, DNN will be further used to predict CKD using some of the data.

Figure 7: Accuracy Results of the five (5) Models

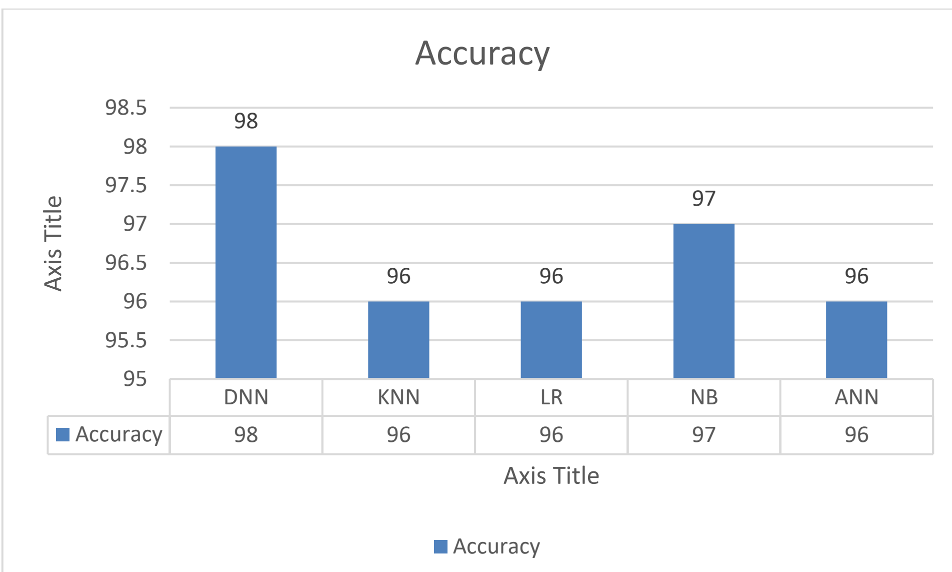

Figure 7 indicates the summary of the accuracy results of the five (5) models used in the prediction of kidney disease, the results shows that DNN have the highest accuracy of 98%, followed by NB: 97%, KNN, LR and ANN: 96% respectively.

Conclusion

In this study, a comparative analysis of five (5) Machine Learning models namely: DNN, KNN, NB, LR and ANN to choose the best technique for the prediction of Chronic Kidney Disease. We analysed and discussed the outcome of these five models, in terms of four (4) performance metrics; accuracy, precision, recall, and F1 Score. Based on the performance metrics of the applied ML techniques, in terms of accuracy, it was revealed as DNN: 98%, KNN: 96%, NB: 97%, LR: 96% and ANN: 96%. The conclusion of the results shows that DNN model can be applied in real world application as a complete system in assisting physicians to input pathological test as inputs and results is provided based on the machine learning so that it can provide efficient prediction and, faster prediction. Future work can be considered by comparing more ML algorithms used for CKD prediction and different diseases dataset can also be considered.

Reference

[1] Alshebly, O. Q., & Ahmed, R. M. (2019). Prediction and Factors Affecting of Chronic Kidney Disease Diagnosis using Artificial Neural Networks Model and Logistic Regression Model. *Iraqi Journal of Statistical Sciences*, 16(28), 140-159.

[2] Arafat, F., Fatema, K., & Islam, S. (2018). Classification of chronic kidney disease (KD) using data mining techniques (Doctoral dissertation, Daffodil International University), 30(2), 3321-3443.

[3] Arasu, S. D., & Thirumalaiselvi, R. (2017). A Prediction of Chronic Kidney Disease Using Feature based Priority Assigning Algorithm. *International Journal of Applied Engi-neering Research*, 12 (20), 9500-9505.

[4] Ayon, S. I., & Islam, M. (2019). Diabetes Prediction: A Deep Learning Approach. *International Journal of Information Engineering & Electronic Business*, 11(2), 21-33.

[5] Barsoum, R. S. (2006). Chronic kidney disease in the developing world. *New England Journal of Medicine*, 354(10), 997-999.

[6] Chahal, A., & Gulia, P. (2019). Machine Learning and Deep Learning. *International Journal of Innovative Technology and Exploring Engineering (IJITEE*), 8(2), 2278-3075.

[7] Chimwayi, K. B., Haris, N., Caytiles, R. D., & Iyengar, N. C. S. (2017). Risk Level Prediction of Chronic Kidney Disease Using Neuro-Fuzzy and Hierarchical Clustering Algorithm (s), 8(12), 112-119.

[8] Chukwuonye, I. I., Ogah, O. S., Anyabolu, E. N., Ohagwu, K. A., Nwabuko, O. C., Onwuchekwa, U., & Oviasu, E. (2018). Prevalence of chronic kidney disease in Nigeria: systematic review of population-based studies. *International journal of nephrology and renovascular disease*, 1(3)11, 165.

[9] Ganggayah, M. D., Taib, N. A., Har, Y. C., Lio, P., & Dhillon, S. K. (2019). Predicting factors for survival of breast cancer patients using machine learning techniques. *BMC medical informatics and decision making*, 19(1), 1-17.

[10] Iliyas, I. I., Saidu, I. R., Dauda, A. B., & Tasiu, S. (2020). Prediction of

Chronic Kidney Disease Using Deep Neural Network. *FUDMA Journal of Sciences*. 4(4), 34-41.

[11] Koh, H. C., & Tan, G. (2011). Data mining applications in healthcare. Journal of healthcare information management, 19(2), 65.

[12] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1), 89-109.

[13] Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. Artificial Neural Networks: *Methodological Advances and Biomedical Applications*. InTech, 6(2)1-18.

[14] Kriplani, H., Patel, B., & Roy, S. (2019). Prediction of Chronic Kidney Diseases Using Deep Artificial Neural Network Technique. *In Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, 23(8), 179-187.

[15] Levey, A. S., Atkins, R., Coresh, J., Cohen, E. P., Collins, A. J., Eckardt, K. U., ... & Powe, N. R. (2007). Chronic kidney disease as a global public health problem: approaches and initiatives–a position statement from Kidney Disease Improving Global Outcomes. *Kidney international*, 72(3), 247-259.

[16] Luyckx, V. A., Tonelli, M., & Stanifer, J. W. (2018). The global burden of kidney disease and the sustainable development goals. Bulletin of the World Health Organization, 96(6), 410-414.

[17] Misir, R., Mitra, M., & Samanta, R. K. (2017). A reduced set of features for chronic kidney disease prediction. *Journal of pathology informatics*. 1(6), 333-340.

[18] Norouzi, J., Yadollahpour, A., Mirbagheri, S. A., Mazdeh, M. M., & Hosseini, S. A. (2016). Predicting renal failure progression in chronic kidney disease using integrated intelligent fuzzy expert system. *Computational and mathematical methods in medicine*, 7(37), 103-110.

[19] Padmanaban, K. A., & Parthiban, G. (2016). Applying machine learning techniques for predicting the risk of chronic kidney disease. *Indian Journal of Science and Technology*, 9(29), 1-6.

[20] Sathya, P. S., & Suresh, K. M. (2018). Chronic Kidney Disease Prediction Using Machine Learning. *International Journal of Computer Science and Information Security (IJC-SIS)*, 16(4), 11-20.

[21] Sharma, S., & Parmar, M. (2020). Heart Diseases Prediction using Deep Learning Neural Network Model. International Journal of Innovative Technology and Exploring Engi-neering (IJITEE), 9(3), 90-102.

[22] Ulasi, I. I., & Ijoma, C. K. (2010). The enormity of chronic kidney disease in Nigeria: the situation in a teaching hospital in South-East Nigeria. *Journal of tropical medicine*, 4(4), 1422-1430.

[23] Ummate, I., Nwankwo, E., & Yusuph, H. (2008), *Journal of Medical Sciences*,2(2), 48-52.