

PREDICTION OF CHRONIC KIDNEY DISEASE USING DEEP NEURAL NETWORK

¹Iliyas Ibrahim Iliyas, ²Saidu Isah Rambo, ³Ali Baba Dauda and ⁴Suleiman Tasiu

¹Department of Computer Science, Nigerian Defence Academy, Nigeria

²Department of Intelligence and Cyber Security, Nigerian Defence Academy, Nigeria

³Department of Mathematical Sciences, University of Maiduguri, Nigeria

⁴Department of Computer Science and Information Technology, Federal University Dutsinma, Nigeria,

Corresponding author's email: iliyasibrahim@nda.edu.ng

ABSTRACT

Deep Neural Network (DNN) is now applied in disease prediction to detect various ailments such as heart disease and diabetes. Another disease that is causing a threat to our health is kidney disease. This disease is becoming prevalent due to substances and elements we intake. Ignoring the kidney malfunction can cause chronic kidney disease leading to death. Frequently, Chronic Kidney Disease (CKD) and its symptoms are mild and gradual, often go unnoticed for years only to be realized of late. We conducted our research on CKD in Bade, a Local Government Area of Yobe State in Nigeria. The area has been a center of attention by medical practitioners due to the prevalence of CKD. Unfortunately, a technical approach in culminating the disease is yet to be attained. We obtained a record of 1200 patients with 10 attributes as our dataset from Bade General Hospital and used the DNN model to predict CKD's absence or presence in the patients. The model produced an accuracy of 98%. Furthermore, we identified and highlighted the Features importance to rank the features used in predicting the CKD. The outcome revealed that two attributes: Creatinine and Bicarbonate, have the highest influence on the CKD prediction.

Keywords: Algorithm, Deep Learning, Medical, Machine Learning, Model

INTRODUCTION

Machine Learning is the scientific field dealing with how machines learn from experience. For many scientists, the term “machine learning” is identical to the term “artificial intelligence”, given that the possibility of learning is the main characteristic of an entity called intelligent in the broadest sense of the word (Joshi & Chawan, 2018). The purpose of machine learning is to construct computer systems that can adapt and learn from their experience.

Chronic Kidney Disease (CKD) is a disorder against proper function regarding kidneys. As kidneys filter our blood whenever kidney disease gets worse, our blood receives wastes at a higher level, which results in sickness. Reduced Glomerular Filtration Rate (GFR) increased urinary albumin excretion, or both are the binding definition terms for CKD. Machine learning to extract useful information can be lent in solving various problems, kidney disease included. Purusothaman & Krishnakumari (2015) indicated that kidney failure falls among several diseases such as heart disease, blindness, etc. resulting in chronic diabetes. Dialysis and transplant are the only methods to keep the kidneys function artificially, and it is also a painful and expensive process. According to Luyckx & Stanifer (2018), kidney disease increased globally from 19 million in 1990 to 33 million in 2013 in 2016, and in 2010 2.62 million people received dialysis worldwide. The need for dialysis is projected to double by 2030.

In Nigeria, the situation is such that CKD represents about 8-10% of hospital admissions (Ulasi & Ijoma, 2010). Therefore, a diagnosis technique is needed to take control or precautions before it becomes late. Various data mining techniques were useful to acquire important information

from medical databases (Chahal & Gulia, 2019). By combining machine learning and statistical analysis, helpful information can be drawn from medical databases.

Machine learning methods coordinate various statistical analyses and databases to extract hidden patterns and relationships from huge and multiple variable data. The accuracy of a given classifier is ensured through testing the model or technique. Moreover, attributes like Specificity, Sensitivity, and Accuracy are common for disease detection (Padmanaban & Parthiban, 2016).

This paper used a Deep Neural Network (DNN) model to predict CKD and evaluate the model's performance by computing the Specificity, Sensitivity, Recall, Receiver Operating Characteristic (ROC) Score, Cohens Kappa, F1 Score, and Precision.

The rest of the paper is organized as follows: Section 2 provides the related work on machine learning and DNN. Section 3 presents how the research is carried out, while section 4 provides the results and the discussion of the research. In section 5, we provided the conclusion of the paper, recommendations, and direction for further work.

RELATED WORK

A number of research has been conducted in machine learning for building models that assist in predicting different types of diseases and health-related problems using different DNN algorithms. This section presents a

review of some of the conducted research in machine learning in Chronic Kidney disease prediction.

Ge et al. (2019) predicted Parkinsons disease severity using DNN using the University of California Irvine (UCI) Parkinson's telemonitoring voice dataset of patients. The studies comprised a biomedical voice measurement of 42 patients with Parkinsons Disease (PD). Severity prediction based on total Unified Parkinsons Disease Rating Scale (UPDRS) accuracy score of 94.4422% and 62.7335% for train and test dataset, respectively, was discovered. The PD severity based on total UPDRS accuracy scored 83.367% and 81.6657% for the train and test dataset.

Ayon & Islam (2019) proposed a strategy for the diagnosis of diabetes using DNN on the PIM Indian Diabetes (PID) dataset from UCI machine learning repository and revealed an accuracy of 98.35%, F1 Score: 98%, and MCC:97% for five-fold cross-validation. Additionally, the accuracy of 97.11%, Sensitivity: 96.25%, and Specificity: 98.80% was obtained for ten-fold cross-validation and indicated that five-fold cross-validation showed better performance.

Shafi et al. (2020) proposed a machine learning-based solution to avoid cleft in the mother's womb with the Deep Learning method and the other four methods on 1000 pregnant female samples from 3 different hospitals Lahore, Punjab. They performed data cleaning, scaling, and feature selection method and compared the accuracy for all the algorithms using Random Forest (RF) algorithm: 85.77%, Decision Tree (DT): 88.14%, K-Nearest Neighbor (KNN): 89.72%, Support vector Machine (SVM): 90.69% and Multilayer perception (MLP): 92.6% and the result indicated that MLP yielded a better accuracy.

Sharma & Parmar (2020) proposed a model for heart disease prediction with the DNN model on heart disease from a UCI dataset of 14 attributes and 303 columns with six (6) different classifiers KNN, SVM, NB, RF, and DNN using Talos optimization, Talos which is an automated POD (Prepare, Optimize, Deploy) pipeline that stability yields state by step prediction results in a wide range of prediction related problems. Their work indicated an accuracy for KNN: 90.16%, Logistic Regression: 82.5%, SVM: 81.97%, NB: 85.25%, and DNN with Talos optimization: 90.78%.

Ahmed & Alshebly (2019) applied Artificial Neural Network (ANN) and Logistic Regression (LR) on 153 cases and 11 attributes of CKD patients. The study used a problem in the domain of medical diagnosis and analyzed each technique efficiency. The outcome revealed that ANNs classifier is better than LR mode with an accuracy of 84.44%, a sensitivity of 84.21, a specificity of 84.61%, and Area Under the Curve (AUC) 84.41%. It was further found that the most critical factors that have a clear impact on chronic kidney disease patients are creatinine and urea.

Chimwayi et al. (2017) applied a neuro-fuzzy algorithm to determine the risk of CKD patients using a dataset of 25 attributes (14 nominal and 11 numeric) describing early stages of CKD in Indians, with an accuracy of 100%,

sensitivity of 100% and specificity of 97%. The authors suggested that the work should be added to the healthcare domain and can be used to provide suggestions in the domain by simplifying the diagnosis and treatment of patients and identifying relationships in the diseases suffered by patients.

Kriplani, Patel & Roy (2019) Studied 224 records of chronic kidney disease available on the UCI machine learning repository named chronic kidney diseases dating back to 2015 and proposed an algorithm. The proposed method is based on a DNN, which predicts the presence or absence of chronic kidney disease with an accuracy of 97%. Compared to other available algorithms, the model built shows better results when implemented using the cross-validation technique by keeping the model safe from overfitting.

Norouzi et al. (2016) presented an Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict CKD's renal failure time frame on a 10-year real clinical data of diagnosed patients. The dataset had 10 attributes. ANFIS model was used for estimating Glomerular Filtration Rate (GFR) at subsequent 6, 12, or 18 months. The model achieved an accuracy of 95%. Urine protein was not used among its parameters. They concluded that the comparisons of the predicted values with the real data showed that the ANFIS model could accurately estimate GFR variations in all sequential periods (Normalized Mean Absolute Error lower than 5%).

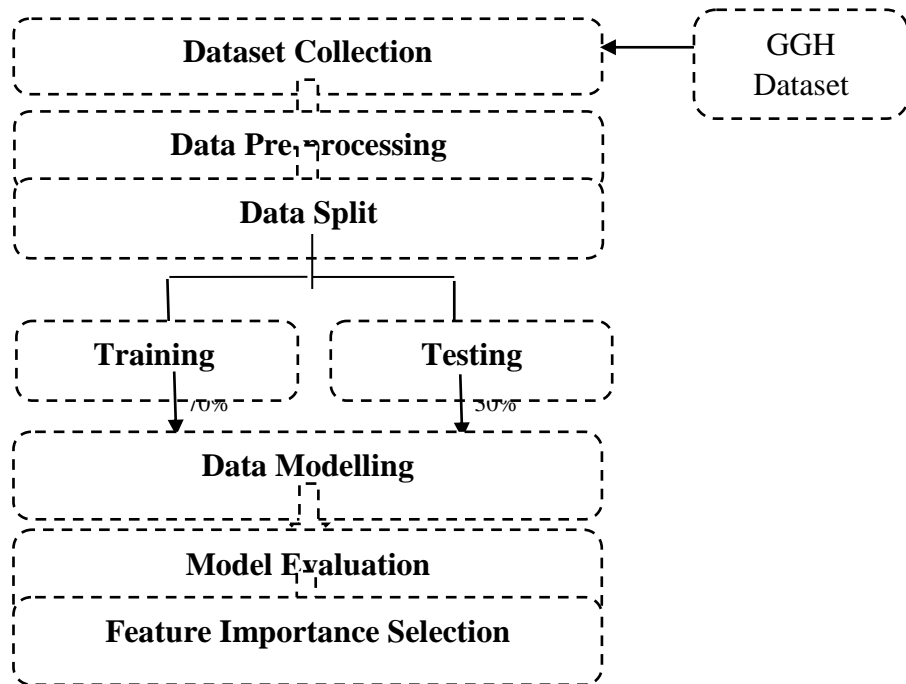
Sathya & Suresh (2018) employed Decision Tree (DT) and Naïve Bayes(NB) as machine learning algorithm to predict CKD using UCI machine repository dataset with 25 attributes and 400 instances, with the use of sample split validation technique an accuracy of 99.25% and 98.75% for DT and NB respectively, showing DT as a better algorithm in terms of predicting the presence and absence of CKD.

Arafat, Fatema & Islam (2018) Studied an automated detection of CKD with clinical data using RF, NB, and DT 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 a higher accuracy of 97.5% and 96% for Naïve Bayes and Decision Tree in predicting CKD. The correlation and importance of each of the attributes on the dataset was explored in their study

Based on the literature presented, the studies performed research on predicting CKD's presence and absence using different algorithms. In this work, we applied DNN to data collected from 1200 patients. Our work focused on using the DNN model to CKD with a split-sample validation technique on a primary dataset from Gashua General Hospital (GGH) Yobe State, Nigeria. To evaluate this model's performance on the dataset, we computed eight (8) performance measures. Also, we performed feature importance selection to determine rankings of the attributes in the dataset that have the highest influence in CKD's prediction.

METHODOLOGY

This methodology depicts the design of the method to be exploited to experiment. It incorporates data collection, data preprocessing, target variable selection, prediction DNN model, and performance evaluation, as shown in Figure 1.



DATA COLLECTION

The patient’s kidney disease record is selected as the source of data for this work. The dataset is collected from General Hospital in Gashua Local Government Area of Yobe State from 2010 to 2019. The dataset contains 1200 patients record with 11 attributes/parameters: 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 Not-Chronic Kidney Disease (NCKD) disease as shown in Table 1.

Table1: Dataset Attributes

Attribute	Attribute Description
Sex	Gender
Age	Age
Sod	Sodium
Pot	Potassium
Chl	Chloride
Bica	Bicarbonate
Urea	Urea
Cre	Creatinine
UA	Uric Acid
Alb	Albumin
Class	{Chronic Kidney Disease, Not Chronic Kidney Disease}

Data Pre-processing

Data Preprocessing represents the most critical task. It involves cleaning, extraction, and transforming data into a suitable format for machine execution. The raw data contains missing information, wrong formats, and invalid information, leading to an error in the prediction. The dataset used had some missing cells replaced using simple imputation with the attribute's mean value. The attribute Sex was converted to numeric values as ‘1’ and ‘0’ for Male and Female, respectively, using one-hand encoding to make it possible for the machine to process since the machine will not understand string values.

Data Modelling

We used a model that has one (1) input, two (2) hidden layers, and one (1) output in the adopted model. The input layer is a set of neurons that take the actual input data for processing. For input and the hidden layer, the “Relu” activation function was used. The output layer has only one output result, either CKD or NCKD. We used the Sigmoid activation function. Stochastic Gradient was used as the optimizer of the model. The experiments are constructed with Python Programming Language. Figure 2 shows the structure of the DNN model adopted for this study.

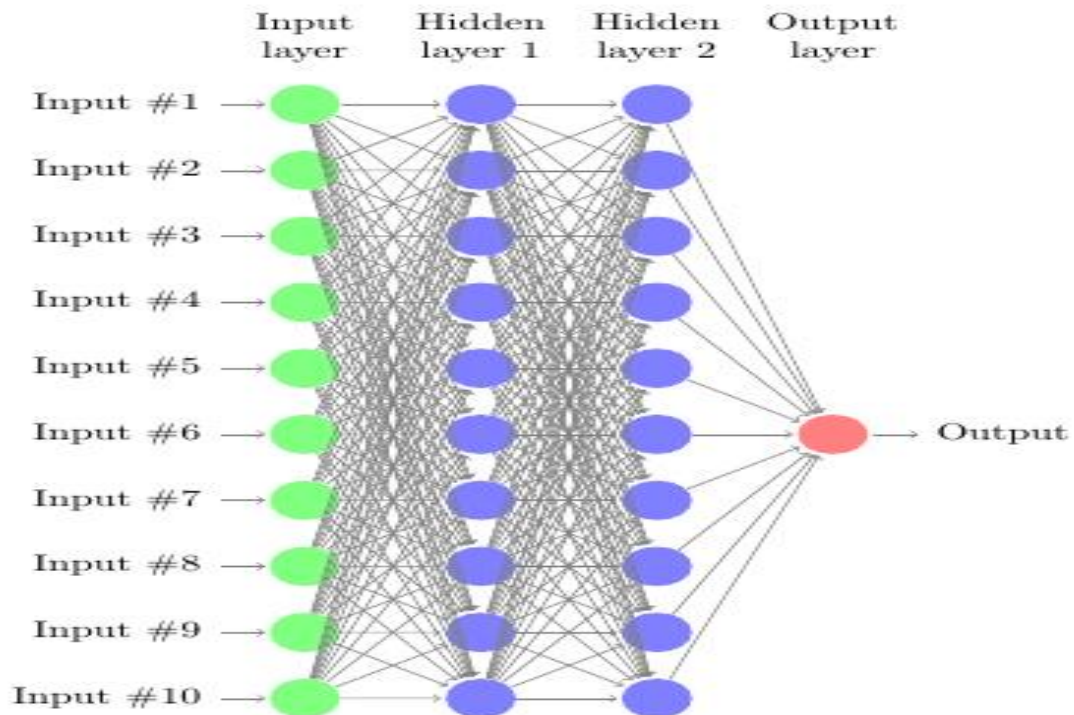


Figure 2: DNN Structure

3.4 Evaluation of the Model

In this work, the performance is measured by accuracy, specificity, sensitivity, Cohen Kappa, precision, F1 score, ROC Score, and recall described as follows.

- Confusion Matrix: confusion matrix indicates the statistical suitability of the model and its compatibility with the dataset; it can also be defined as a table layout that is used explicitly for visualization of algorithm performance. Table 2 shows a summary of the confusion matrix.

Table2: Confusion Matrix

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

- Accuracy is used to identify the number of correctly predicted data points out of all data points. It is defined as the number of all correct predictions made divided by the total number of predictions. It is expressed in equation 3.1;
- Sensitivity- (Recall or True Positive Rate): it is defined as the proportion of actual positive cases that got predicted as positive. It is a ratio of true positive to the sum of a true positive and false negative. In medical diagnosis, test sensitivity (Recall) is a test's ability to correctly identify those with the disease. It is expressed in equation 3.2;

- Specificity- (True Negative Rate): it is defined as the proportion of actual negative that got predicted as the negative. It is calculated as the number of correct negative predictions divided by the total number of negatives. It is also called true negative rate. It is expressed in equation 3.3;
- Cohen Kappa: This is a classifier performance measure between two sets of classified data. Kappa result values are between 0 to 1. The results become meaningful with increasing values of kappa. It measures how closely instances classified by the machine learning classifier matched the data labeled as the truth. It is expressed in equation 3.4;
- Precision: It is defined as the fraction of relevant instances among the retrieved instances. This is given as the correlation number between the correctly classified modules to entire classified fault-prone modules. It is expressed in equation 3.5.
- Recall/ Sensitivity: Recall is a metric that quantifies the number of correct positive predictions made from all positive predictions that could have been made. It is expressed in equation 3.6.
- F1 Score: This is the harmonic mean between precision and recall. The range for f1-score is from 0 to 1. 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. The expression of F1-score is as follows;

RESULTS AND DISCUSSION

This section provides the results and discussion of the DNN prediction on the CKD.

Results

This section provides the results of the DNN algorithm produced based on the dataset.

Table 3: Confusion Matrix for Deep Neural Network

Classification		Prediction	
		Absence	Presence
Observation	Absence	92	2
	Presence	1	85

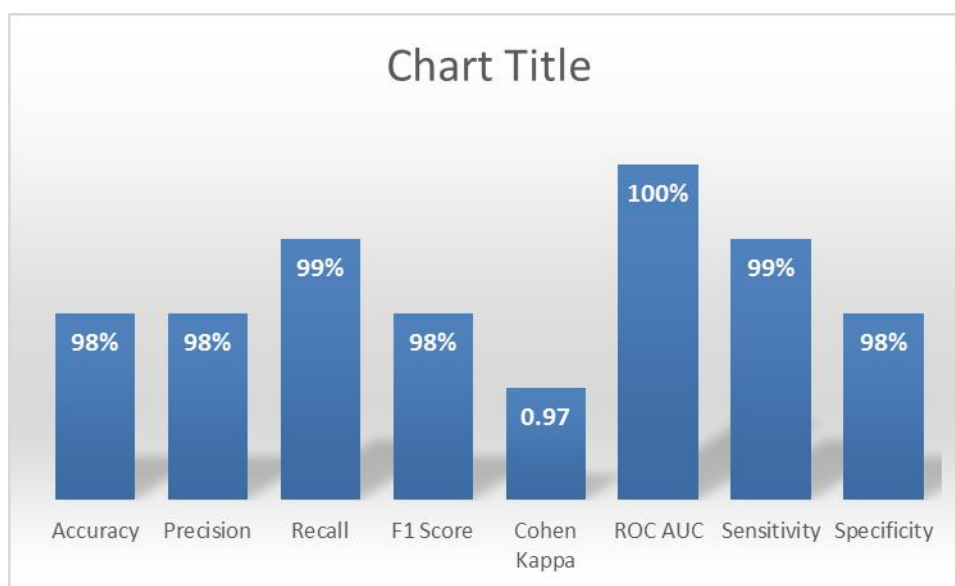


Figure 3: Performance evaluation for Deep Neural Networks Model

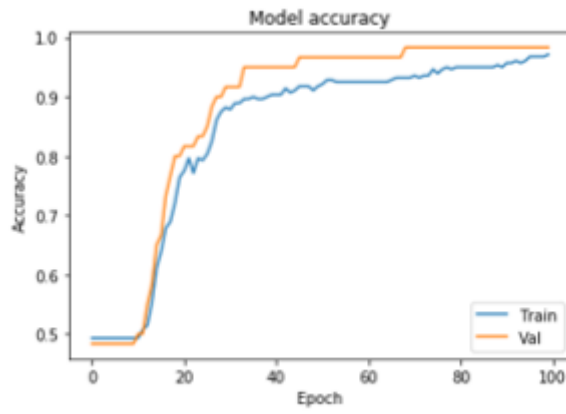


Figure 4: Graph of training against Validation Accuracy

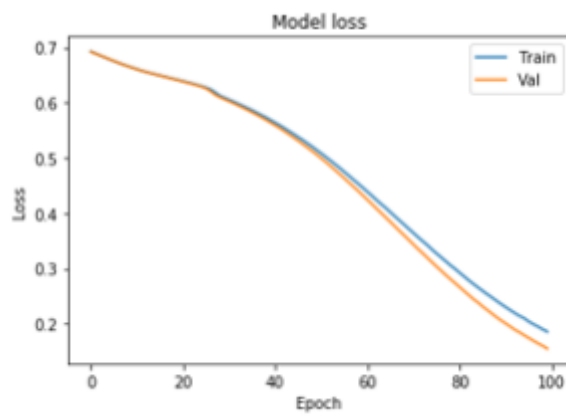


Figure 5: Graph of Training against Validation Loss

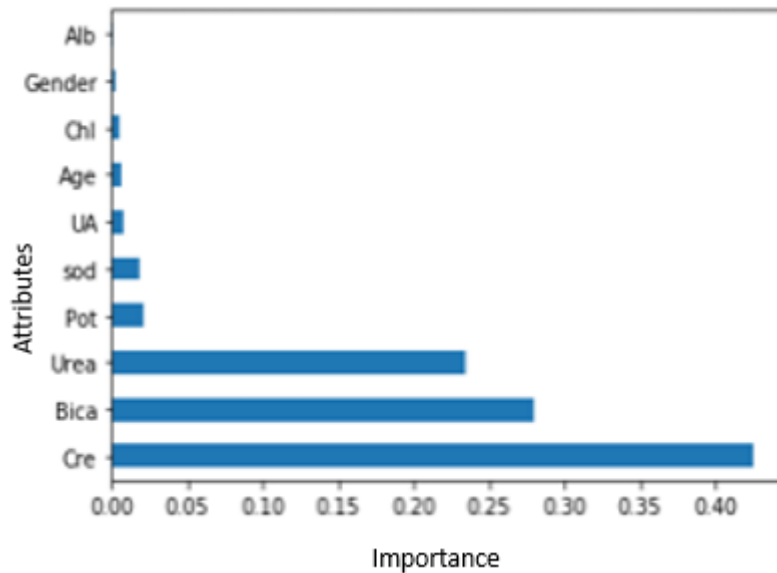


Figure 6: Feature Importance

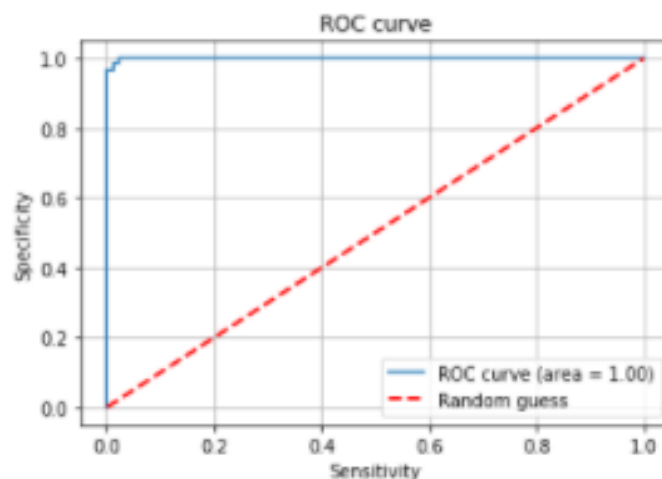


Figure 7: ROC Curve

DISCUSSION

We divided our dataset into 70% and 30% for testing and training respectively, Table 3 shows the confusion matrix provided a visualization summary of the DNN model on the dataset. The model has succeeded in classifying 92 samples correctly, which are NCKD and 85 samples correctly, CKD, but 1 sample, which is CKD, was classified wrongly, and 2 samples that are NCKD were classified wrongly. Totaling to 177 samples classified correctly, indicating that the DNN model performs well on our dataset since only 3 samples out of the sample selected for testing was not classified from the 180 samples for the testing set.

Figure 3 represented the model's performance evaluation and indicated that the model had succeeded in predicting kidney patients' records into CKD and NCKD with 98% overall prediction accuracy when tested with unseen data. Based on this accuracy of 98%, the model indicated that it is hyper-intelligent to predict CKD and NCKD seamlessly. Other metrics used and evaluated the performance produced Precision :98%, Recall: 99%, F1 Score: 98%, Cohens Kappa: 0.97%, ROC Score: 100%, Sensitivity:99% and Specificity:98%.

Figure 5 shows the accuracy loss of both the training sets and validation sets keeps reducing at the same level. Indicating no overfitting since there is an improvement because iterates, the loss accuracy decreases in both the training and testing sets with almost the same level of declination. Figure 4 shows our model accuracy level as iterated in both the training set and validation set. Indicating accuracy improvement of our model for both training and testing set. This accuracy has almost the same advancement level, with the validation set sometimes surpassing the training set, indicating our model's effectiveness and no overfitting.

Figure 6 shows the ranking of features in the dataset that influence CKD prediction from the 10 attributes. Creatinine and Bicarbonate have the highest ranking, followed by Urea, Potassium, Sodium, Uric Acid, Age, Chloride, Gender, and Albumin. It was revealed that Creatine and Bicarbonate have much influence during the prediction.

Figure 7 highlighted the plotted graph of the ROC Curve of the DNN model used, and the curve indicated to be closer

to the top left corner of the graph, showing the better performance of the model.

SUMMARY, CONCLUSION, AND FURTHER RESEARCH

This section summarizes the paper, the conclusion by identifying the performance of the DNN and the attributes that were found too pronounced. The section also enumerates some research that can be done in the future to provide effective performance.

SUMMARY

This work collected and trained the DNN model to predict the absence and presence of Chronic Kidney Disease with an accuracy of 98%, Precision: 98%, Recall: 99%, F1 Score: 98%, Cohens Kappa: 0.97%, ROC Score: 100%, Sensitivity:99%, and Specificity:98% was used to evaluate the performance of the adopted model and provided a summary of attributes influencing CKD prediction.

CONCLUSION

This research's main goal is to use the DNN model to predict kidney disease to a high degree of accuracy. We succeeded in classifying the kidney disease dataset into CKD and NCKD with 98% overall accuracy when the model was tested with a set of data not used during the training process. The adopted DNN model proved to be efficient and suitable for the prediction of kidney disease. The study also highlighted the importance of the features used in the prediction of kidney disease. This revealed that from the 10 attributes, Creatinine and Bicarbonate are the attributes with the highest influence on CKD as opposed to the findings by Ahmed & Alshebly (2019) that found that Creatinine and Urea have the highest influence on CKD prediction.

FURTHER RESEARCH

Incorporating more data to have large dataset will provide better accuracy and efficiency; hence more dataset is needed to accommodate enough samples. Having enough samples will make the prediction wider to capture and identify regions and areas with CKD vulnerability. This

paper used the DNN model only. Having other models to compare techniques may provide better performance.

REFERENCE

- Ahmed, R. M., & Alshebly, O. Q. (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.
- Padmanaban, K. R., & 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
- Arafat, F., Fatema, K., & Islam, S. (2018). Classification of chronic kidney disease (CKD) using data mining techniques (Doctoral dissertation, Daffodil International University).
- Ayon, S. I., & Islam, M. (2019). Diabetes Prediction: A Deep Learning Approach. *International Journal of Information Engineering & Electronic Business*, 11(2).
- Chahal, A., & Gulia, P. (2019). Machine learning and deep learning. *International Journal of Innovative Technology and Exploring Engineering*, 8(2), 2278-3075.
- Chimwayi, K. B., Haris, N., Caytiles, R. D., & Iyengar, N. C. S. N. (2017). Risk Level Prediction of Chronic Kidney Disease Using Neuro-Fuzzy and Hierarchical Clustering Algorithm (s). *International Journal of Multimedia and Ubiquitous Engineering*, 12(8), 23-36
- Chollet, F. (2018). Deep Learning Using Python. In 2018 21st *International Conference on Information Fusion, FUSION*.
- Ge, Y., Wang, Q., Wang, L., Wu, H., Peng, C., Wang, J. & Yi, Y. (2019). Predicting post-stroke pneumonia using deep neural network approaches. *International Journal of Medical Informatics*, 132, 103986.
- Joshi, T. N., & Chawan, P. M. (2018). Logistic Regression and SVM Based Diabetes Prediction System. *International Journal For Technological Research In Engineering*, 11(5), 2347-4718.
- 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* (179-187).
- 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), 414.
- 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*.
- Purusothaman, G., & Krishnakumari, P. (2015). A Survey of Data Mining Techniques on risk prediction: Heart disease. *Indian Journal of Science and Technology*, 8(12), 1.
- Sathya, P. S., & Suresh, K. M. (2018). Chronic Kidney Disease Prediction Using Machine Learning. *International Journal of Computer Science and Information Security (IJCSIS)*, 16(4).
- Shafi, N., Bukhari, F., Iqbal, W., Alm Mustafa, K. M., Asif, M., & Nawaz, Z. (2020). Cleft prediction before birth using deep neural network. *Health Informatics Journal*, 1(18) 1460458220911789.
- Sharma, S., & Parmar, M. (2020). Heart Diseases Prediction using Deep Learning Neural Network Model. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(3).
- 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*.



©2020 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.