Journal of Electrical Engineering, Electronics, Control and Computer Science – JEEECCS, Volume 9, Issue 32, pages 21-28, 2023

An Optimized Diagnostic Model for Chronic Kidney Disease

Osaseri Roseline Oghogho
Department of Computer Science, Faculty of Physical Sciences
University of Benin, P.M.B
Benin, Nigeria
roseline.osaseri@uniben.edu

Amadin Frank Iwebuke
Department of Computer Science, Faculty of Physical Sciences
University of Benin, P.M.B
Benin, Nigeria
frankamadin@uniben.edu

***Abstract* –Chronic Kidney Disease (CKD) is the gradual loss of renal function over a period of time. It is a major global health problem and a leading cause of death especially in developing nations due to the prevalence of risk factors such as hypertension, cardiovascular disease and diabetes. Chronic kidney disease is considered a "silent killer" because there are few symptoms in its early stages. However, different machine learning techniques have been employed in diagnosis of CKD. Amongst these techniques that have been employed in the diagnosis of CKD, Adaptive Neuro Fuzzy Inference System (ANFIS) outperformed other existing models. Thus, this study improved on ANFIS model by presenting an optimized diagnostic (AOD) model for diagnosis of CKD. The Model was designed using genetic algorithm for the optimization of clinical decision variables and a hybrid training algorithm consisting of supervised hybrid learning algorithm (Least Square Estimator and conjugate gradient descent) was used to train the model. The result of the research was subjected to convergence rate and optimum performance criteria (sensitivity, specificity, accuracy, type I error, type II error, type I error rate and type II error rate) and compared with an existing ANFIS model. Results for the research showed that AOD model converged at the 12th epoch with a minimum error of 0.7333 and a sensitivity of 100%, specificity of 89%, accuracy of 95%, type I error of 6, type II error of 0, type I error rate of 0 and type II error rate of 0.0857 as against the ANFIS model which converged at the 72nd epoch with minimum error of 0.7333 and a sensitivity of 100%, specificity of 77 %, accuracy of 88%, type I error of 15, type II error 0, type I error rate of 0 and type II error rate of 0.2143. This research shows that AOD model resulted in better classification accuracy and also takes less time to converge than ANFIS model. -electronic document is a "live" template. The various components of your paper [title, text, heads, etc.] are already defined on the style sheet, as illustrated by the portions given in this document.**

Keywords-chronic kidney disease, Optimization, ANFIS, Diagnosis, Model, machine learning

I. INTRODUCTION

The kidneys are bean shaped organs of the urinary system, approximately the size of a fist that serves important roles in human. The kidneys perform an essential job by filtering the blood and returning it to the circulatory system [27]. In a day the human kidney filter about a hundred and twenty to one hundred and fifty quarts of blood to produce about one to two quarts of urine, which composed of wastes and extra fluid [28]. Kidney diseases also called renal diseases refers to many kinds of diseases which are disorders that affects the kidneys or any damage that reduces the function of the kidney. When urinary organ performance declines, the competency of the kidneys to filter blood is expeditiously reduced. There are several causes of chronic kidney disease, such as diabetes mellitus and high blood pressure. People could also be born with abnormalities which will have an effect on their kidneys [27].

Chronic Kidney Disease (CKD) is identified as urinary organ damage in glomerular filtration rate, (GFR) 60 ml/min/1.73 m2 for $\geqslant$ 3 months. Hypertension and diabetes mellitus are two major causes of CKD. High blood pressure is one among the driving causes of CKD because of the pernicious impact that enlarged blood Pressure hose on organ vasculature. Long-term, uncontrolled, high Blood Pressure leads to high intra glomerular pressure, disabling glomerular filtration and diminished renal ability interferes with the kidneys' function to maintain fluid and solution equilibrium [13].

The ability to concentrate urine declines early and is followed by decreases in ability to excrete phosphate, acid, and potassium [12], [19]. The final stage of CKD is called; end-stage renal disease. At this stage, the kidneys will no longer be able to remove the waste produced from the body as excess fluid.

Early diagnosis of CKD at presentation in primary health care facilities and prompt referral to nephrologists for adequate management can reduce the morbidity and death rate associated with CKD. Thus, the application of machine learning approach by physicians, primarily at the first health care level in diagnosing of CKD can accelerate the referral process of cases to the nephrologists which ultimately will help to curb the menace of CKD. However, existing work has employed different machine learning techniques in the diagnosis of renal disease, though it had perpetually been a tricky task to spot the simplest technique for any diagnosis, because different techniques have their own limitation in view of performance and convergence time [16], [1], [6].

Existing models have gained various degrees of success, though not without limitation, such limitations include feature selection (attributes optimization) and slow training (convergence rate). A good understanding of a model limitations and drawbacks is vital in machine learning [10].

The aim of this research is to propose An Optimized Diagnostic (AOD) model for CKD; the model will make use of optimization technique for feature selection of clinical parameters and develop a training algorithm for faster convergence and improved diagnostic performance.

An overview of CKD was carried out as well as related literature on techniques in CKD in section 1 and 2. In section 3, the methodology of AOD model development was discussed, the genetic algorithm that was used for the feature selection were presented, and the data were preprocessed for the training. Section 4 was the presentation of result, the various evaluation and tests that were carried out on the AOD model before and after optimization of the features of the dataset were discussed.

II. Ease of Use Background

Reference [1] proposed a neural network algorithm model for kidney stone diagnosis; the model was compared with Back Propagation Algorithm (BPA), Radial Basis Function (RBF) and non-linear classifier Support Vector Machine (SVM). The result shows that BPA is the best model for training neural network for kidney stone diagnosis. Reference [7] proposed a clinical decision support system for the diagnosis and management of chronic renal failure using ANN, Naïve Bayes and Decision tree. The obtained result shows that the decision tree algorithm was the most accurate classifier.

Reference [21] analyzed ANN, Decision tree, and Logic regression supervised machine learning algorithm. The algorithms were used for kidney dialysis. Reference [3] carried out a study on the diagnosis of renal failure disease using ANFIS; the performance was based on classification accuracy, sensitivity and specificity. The study compared Support Vector Machine (SVM), ANN and ANFIS using the same training and testing datasets.

The experimental result shows that ANFIS classification accuracy was better than SVM and ANN. Reference [21] analyzed ANN, Decision tree, and Logic regression supervised machine learning algorithm on kidney dialysis. Tanagara data mining tools were used for classification. The experimental result showed that ANN out performed decision tree and logical regression. Reference [8] used two data mining algorithms; ANN and SVM to predict patients' death and their need for dialysis.

The experimental result showed that ANN out performed SVM in predicting patients' deaths with a classification accuracy of 82.7%, sensitivity of 83.4% and a specificity of 82.3%. While in predicting patients' need for dialysis, ANN achieved classification accuracy of 82.3%, sensitivity of 83.1% and a specificity of 84.3%. The result from this study suggested that ANN can provide a good kidney disease prognosis, though it takes a long time in terms of training. Reference [32] analyzed two classification algorithms, ANN and SVM for prediction of kidney disease.

The comparative analysis of the two-classification algorithm shows that ANN is a better classifier. Reference [29] proposed three classification algorithms; radial basis function network, multilayer perceptron, and logistic regression for prediction of CKD. Comparative analysis was carried on the three classifiers. The result of the experiment showed that multilayer perceptron has a better performance than the other classifiers.

Reference [33] presented two classification algorithms Naives Bayes and SVM for prediction of kidney disease. The algorithm performance was determined based on classification accuracy and execution time. The result showed that SVM achieved an increase in classification performance than naïve Bayes. However, SVM has better classification accuracy than naïve bayes but it takes long time to execute. Reference [28] carried out a study on the prediction of chronic kidney disease. The proposed model was assessed using four different measurement criteria such as kappa, accuracy, sensitivity and specificity. Radial Basis Function has a better accuracy for predicting chronic kidney disease, with an accuracy of 85.3%. Reference [7] presented three learning algorithms on a set of medical data with the objective to predict kidney disease by using support vector machine, Decision Tree (C4.5), and Bayesian Network (BN). The goal was to compare different classification models and define the most efficient one. When implemented on WEKA software; in accuracy C4.5 scored 63%, followed by SVM 60.25% and Bayesian Network 57.5%. Reference [26] proposed ANFIS model for prediction of renal failure progression of chronic kidney disease.

Researchers in soft computing have employed different classification techniques in diagnosis of CKD. From the reviewed literature the existing classification techniques have their peculiar drawbacks; such drawbacks include inability to learn, high computational cost, feature selection (attributes optimization) and learning problem (slow training). As with any model, a proper knowledge of the challenge and drawbacks of such model is an important fact in the machine learning process [10]. A common learning problem for neural networks is intractably which makes training neural networks a computationally hard problem. A promising approach for breaking this problem is to test different model and if fast training algorithms can be designed [17]. However, there are fast training algorithms for neural networks offering better and faster convergence than existing method that is slow in training. By utilizing a technique of nonlinear optimization, it is possible to accelerate the learning method [25].

III. Methodology

There are two types of Machine learning algorithms; the supervised learning and unsupervised

learning algorithms. Supervised learning algorithm is an important research topic in machine learning [30], [35]. Enormous research interest is been devoted towards machine learning algorithm especially for solving classification problem [22], [35]. This research study explores the application of machine learning technique in prediction of CKD using a supervised learning classification algorithm.

A machine learning algorithm approach was employed in training AOD model, and genetic algorithm for the optimization of clinical parameter in the diagnosis of Chronic Kidney Disease. This study developed a hybrid learning algorithm that will improve on existing convergence time and enhance the classification performance of the existing models. Adaptive Neural Fuzzy Inference (ANFIS) model was adapted as our architectural framework in developing AOD model. ANFIS is an artificial Neuro-Fuzzy Inference Systems belonging to a class of adaptive networks that are functionally equivalent to fuzzy inference systems trained with a hybrid learning algorithm and least-squares method to identify the consequent parameters within the forward pass and gradient descent to propagate error in backward pass to update the premise parameters. In order to improve on ANFIS architectural framework, we employed the use of a hybrid learning algorithm that consist of the Least Square Estimator (LSE) and conjugate gradient descent approach which will improve the convergence speed during training of the network [9]. Also, Genetic Algorithm (GA) was employed to optimize the parameters in order to obtain an optimal search for the best parameters [4].

A. Dataset

The dataset for this study was collected from the University of California, Irvine (UCI) Centre for machine learning repository (http//archive.ics.uci.edu/ml/datasets/). The AOD model contains 400 real world CKD datasets made up of 24 attributes having 11 numerical, 13 nominal and 1 class. Table 1 shows the datasets field description.

TABLE I. DATASETS FIELD DESCRIPTION

S/N	FEATURES	LABELS	UNITS	CATEGORY
1	Albumin	al		Nominal
2	Anemia	ane		Nominal
3	Appetite	appet		Nominal
4	Bacteria	ba		Nominal
5	Blood glucose random	bgr	Mgs/dl	Numerical
6	Blood pressure	bp	mm/Hg	Numerical
7	Blood urea	bu	mgs/dl	Numerical
8	Coronary artery disease	cad		Nominal
9	Diabetes mellitus	dm		Nominal
10	Hemoglobin	hemo	gms	Numerical
11	Hypertension	htn		Nominal
12	Pus cell	pc		Nominal
13	Pus cell clumps	pcc		Nominal
14	Packed cell volume	pcv		Numerical
15	Pedal edema	pe		Nominal
16	Potassium	pot	MEq/L	Numerical
17	Red blood cells	rbc		Nominal
18	Red blood cell count	rbcc	Millions/cmm	Numerical
19	Serum creatinine	sc	mgs/dl	Numerical
20	Specific gravity	sg		Nominal
21	Sodium	sod	MEq/L	Numerical
22	Sugar	su		Nominal
23	White blood cell count	wbcc	Cells/cumm	Numerical
24	Age	age	yrs	Numerical

B. Data Preprocessing

The raw dataset from UCI consists of missing values, categorical values and numerical values. The imputation approach of handling missing values was employed in this study [34], [14]. The categorical features in the dataset were replaced with 1 and -1, were 1 indicate present, normal, yes and good while -1 indicate not present, abnormal, no and poor.

The next stage was the normalization of the dataset; normalization is a "scaling down" transformation of the features. The datasets were scaled to the range of (0, 1) using min – max normalization equation as captured in equation (1).

$$x_{ni} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \qquad (1)$$

Where:
x_i = the real-world CKD value
X_{ni} = the scaled input value of the real-world CKD value X_i
X_{min} and X_{max} are the minimum and maximum values of the un-scaled dataset.

C. Feature Selection

Feature selection decreases the training time, enhanced generalization and also improves the prediction accuracy in classification problem [23],[15],[5]. Genetic Algorithm (GA) optimization technique was used to decrease the computational complexity and increased the total classification accuracy of AOD model. Our choice of using GA was influenced by its exploration effectiveness during the search of the space of possible solutions, GA does not assess solutions one by one, but evaluate a set of solutions simultaneously, secondly, it does not require assumptions about the interactions between features, and finally GA does not get stuck in local minima [21]. Hence GA was employed for selection of attributes (Feature selection).

The fitness of each chromosome (feature) was ranked based on percentage of relevance of individual feature to the class prediction. The chromosome population was ranked using algorithm (1) while the fitness evaluation of the chromosomes was carried out using fitness function as captured in equation (2) and algorithm 2. The Fitness is used to ascertain how well the chromosome is able to survive [20].

Algorithm 1: feature ranking algorithm.

```
int ran;
if (val> max * 0.8)
ran = 5;
else if (val> max * 0.6 &&val< max * 0.8)
ran = 4;
else if (val> max * 0.4 &&val< max * 0.6)
ran = 3;
else if (val> max * 0.2 &&val< max * 0.4)
ran = 2;
else if (val> max * 0.1 &&val< max * 0.2)
ran = 1;
else
ran = 0;
```

$$Fitness = \sum_{0}^{n} R_x \quad (2)$$

R_X = Rank of Chromosome position x

n = Number of parameters

D. Classification Algorithm

The proposed AOD model is an adaptive network that uses a supervised learning algorithm, which has a function similar to the model of Takagi–Sugeno fuzzy inference system. The network training is a continuous update of the network parameters. A hybrid algorithm was used for the training of these parameters in order to improve on the slow convergence nature that existed in the gradient decent backpropagation algorithm used in training parameters in ANFIS.

The hybrid learning consists of least square estimator and conjugate gradient algorithm. The least square technique acts as a forward pass to identify the consequent parameters in the fourth layer. The conjugate gradient decent acts as the backward pass and are utilized to fine – tune the premise parameters equivalent to the fuzzy set in the input domain. The output generated from the network called the actual output is compared with the required output also called anticipated output; the error that happened amid the comparison between the required outputs with the actual output is propagated back to the first layer. The foremost objective of the adaptive learning framework is to decrease errors that happen in the network. The AOD model network consists of five layers as represented in figure 1.

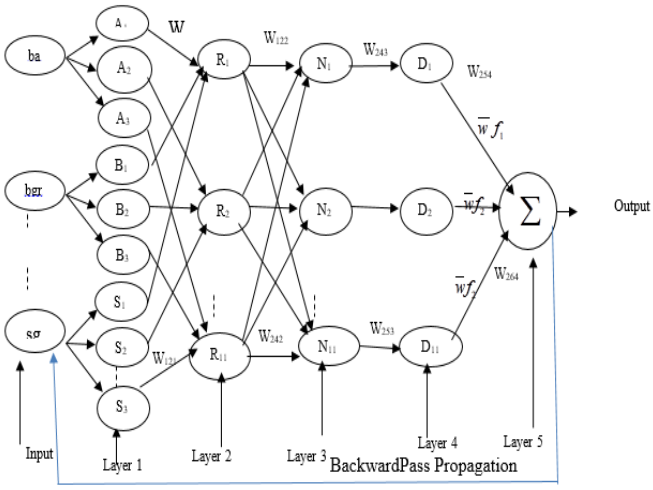

Figure 1. AOD network.

IV. RESULT

The goal of any classification algorithm is to achieve optimum performance and better convergence. CKD dataset from UCI repositories was tested with AOD and ANFIS classifier model, after feature selection the initial twenty-four (24) features were reduced to eleven (11) features as shown in table 2, both models were trained with 280 CKD dataset using the optimized and un-optimized features.

The models were trained (AOD and ANFIS) before and after optimization; after several iterations AOD model after optimization converged at 12th epoch, with a minimum error of 0.7333 while ANFIS model after optimization converged at 72nd epoch with least error of 0.7333.

The result of convergence for the AOD model and ANFIS model after optimization are shown in figure 2 and 3 respectively. AOD and ANFIS model before optimization were also subjected to the same training. AOD model converged at the 22nd epoch while ANFIS model converged at the 144th epoch. The results are represented in figure 4 and 5 respectively.

TABLE II. DATASETS FIELD DESCRIPTION

S/N	FEATURES	LABELS	UNITS	CATEGORY
1	Bacteria	ba		Nominal
3	Blood glucose random	bgr	Mgs/dl	Numerical
4	Blood pressure	bp	mm/Hg	Numerical
5	Blood urea	bu	mgs/dl	Numerical
6	Diabetes mellitus	dm		Nominal
7	Hemoglobin	hemo	gms	Numerical
8	Hypertension	htn		Nominal
9	Pus cell clumps	pcc		Nominal
10	Packed cell volume	pcv		Numerical
11	Serum creatinine	sc	mgs/dl	Numerical

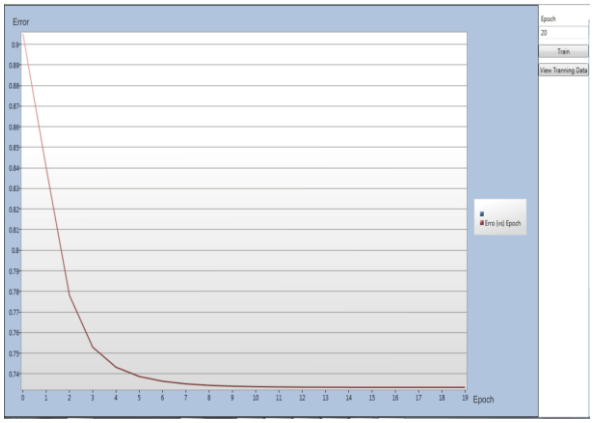

Figure 2. Training error of AOD model using optimized features.

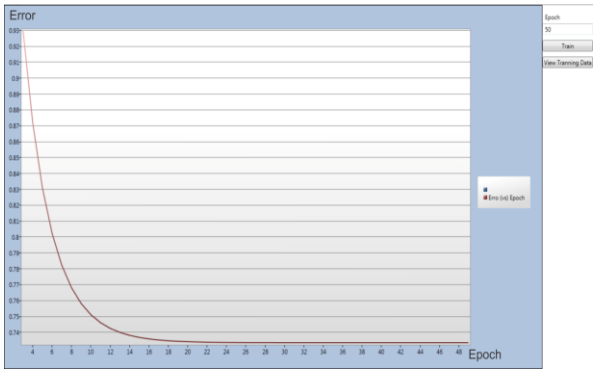

Figure 3. Training error of ANFIS model using optimized features.

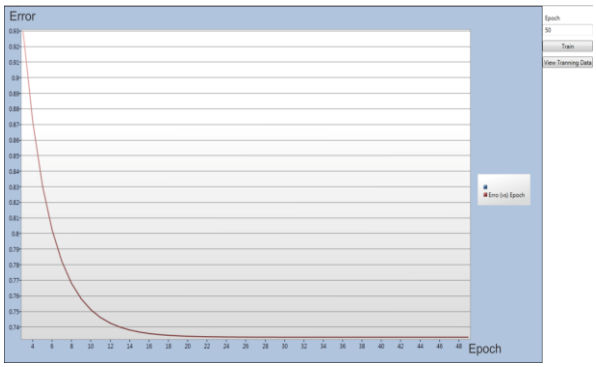

Figure 4. Training error of AOD model using the un-optimized features.

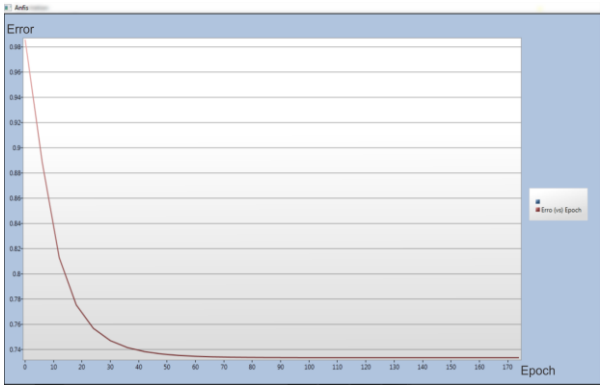

Figure 5. Training error of ANFIS model using the un-optimized features.

AOD and ANFIS models were tested for optimum performance using both optimized and the un-optimized features. In other to determine the better performing classifier in predicting CKD, the following performance measurements were used; sensitivity, specificity, accuracy, type 1 error, type 2 error, type 1 error rate and type 2 error rates. One hundred and twenty (120) cases of CKD were used to test AOD and ANFIS models. The results for AOD and ANFIS models are captured in figures 6- 9.

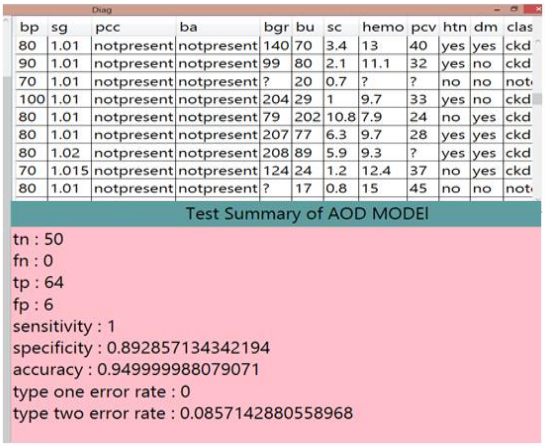

Figure 6. Performance of AOD model Using Optimized Features.

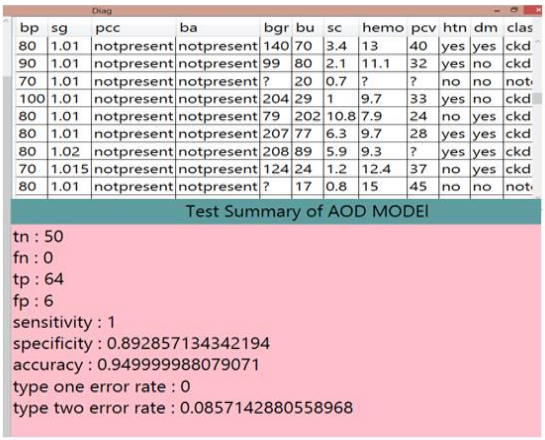

Figure 7. Performance of ANFIS Model Using Optimized Features.

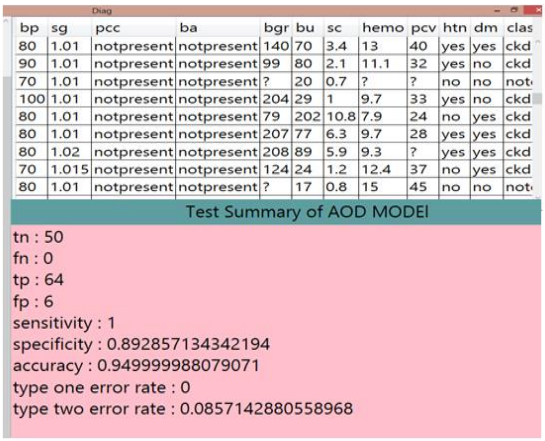

Figure 8. Performance of AOD Model Using Un-Optimized Features.

	pcc	ba	bgr	bu	sc	hemo	pcv	cad	ans	clas
25	notpresent	notpresent	140	10	1.2	15	48	no	no	ckd
25	notpresent	notpresent	70	36	1	17	52	no	no	ckd
25	notpresent	notpresent	82	49	0.6	15.9	46	no	no	ckd
25	notpresent	notpresent	119	17	1.2	15.4	42	no	no	ckd
25	notpresent	notpresent	99	38	0.8	13	49	no	no	ckd
25	notpresent	notpresent	121	27	1.2	13.6	52	no	no	ckd
25	notpresent	notpresent	131	10	0.5	14.5	41	no	no	ckd
2	notpresent	notpresent	91	36	0.7	14	46	no	no	ckd
2	notpresent	notpresent	98	20	0.5	13.9	44	no	no	ckd
2	notpresent	notpresent	104	31	1.2	16.1	45	no	no	ckd

Diagnostic Result

tn : 0
fp : 80
tp : 39
fn : 1
sensitivity : 0.32773110 2705002
specificity : 0
accuracy : 0.324999988079071
type one error rate : 0.672268927097321
type two error rate : 0.025000000372529

Figure 9. Performance of ANFIS Model Using Un-Optimized Features.

V. DISCUSSION

The performance of AOD and ANFIS models were tested on convergence and optimum performance using; accuracy, sensitivity, specificity, type 1 error, type 2 errors, type 1 error rate and type 2 error rates. The convergence behavior of the classifiers as shown in figure 10, we observed that AOD optimized features and ANFIS optimized features models converged at 12th epoch and 72nd respectively while AOD un-optimized feature converged at the 22nd epoch and ANFIS un-optimized feature converged at the 144th epoch.

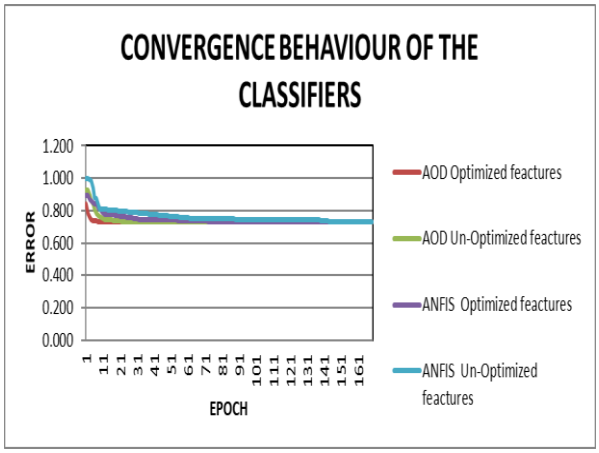

Figure 10. Convergence Behavior of Classifiers.

This infer that fast convergence of models with optimized features is due to the fewer features that were used in the training which further shows that the presence of irrelevant information in data set reduces the speed and quality of learning. Feature selection technique reduces the training time, enhanced generalization and also improves the prediction accuracy in classification problem [23],[15],[5].

The confusion matrix in table 3 is used to summarize the performance of AOD and ANFIS model, which also gave insight about the errors and the type of errors that the models made. where TP is the number of true positives, which means, some cases with positive class (ckd) is correctly classified as positive; FN, the number of false negatives, which means, some cases with the positive class is classified as negative; TN, the number of true negatives, which means, some cases with the negative class (notckd) is correctly classified as negative; and FP, the number of false positives, which means, some cases with the negative class is classified as positive.

TABLE III. DATASETS FIELD DESCRIPTION

CLASSIFIER	CONFUSION MATRIX		TYPE I ERROR RATE	TYPE II ERROR RATE	SENSITIVITY (%)	SPECIFICITY (%)	ACCURACY (%)
AOD MODEL (before feature selection)	**69 (TP)**	**50 (FP)**	**0.42**	**0.0143**	**58**	**0**	**57.49**
	0 (TN)	**1 (FN)**					
AOD MODEL (after feature selection)	**64 (TP)**	**6 (FP)**	**0**	**0.0857**	**100**	**89.28**	**95.09**
	50 (TN)	**0 (FN)**					
ANFIS MODEL (before feature selection)	**39(TP)**	**80(FP)**	**0.67**	**0.025**	**33**	**0**	**32.49**
	0(TN)	**1(FN)**					
ANFIS MODEL (after feature selection)	**55 (TP)**	**15 (FP)**	**0**	**0.2143**	**100**	**76.92**	**87.50**
	50 (TN)	**0 (FN)**					

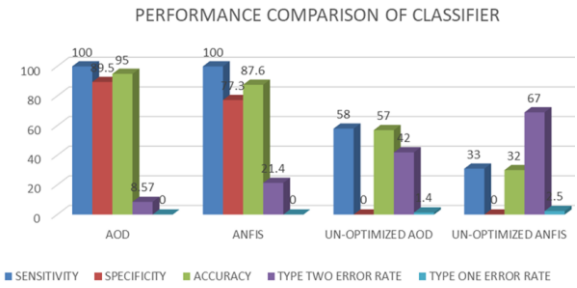

Figure 11. A graphical comparison of the models.

The graphical performance of the models is represented in figure 11. AOD model using optimized features gave the following results; Sensitivity 100%, specificity 89%, accuracy 95%, type 1 error rate 0, type 2 rate 0.085 while ANFIS model with optimized features results are; Sensitivity 100%, specificity 77%, accuracy 86%, type 1 error rate 0, type 2 rate 0.214. When un-optimized features were used to test AOD and ANFIS models the following results were obtained; AOD are Sensitivity 58%, specificity 100%, accuracy 58%, type 1 error rate 0.42, and type 2 rates 0.014 while ANFIS model recorded Sensitivity 33%, specificity 0%, accuracy 33%, type I error rate 0.67, and type II rate 0.025.

We observed from the confusion matrix table that AOD model after feature selection recorded one hundred and fourteen (114) instances that were correctly classified with six (6) instances wrongly classified and the proportion of actual positive cases that were identified by the model were high. ANFIS model after feature selection recorded one hundred and five (105) instances that were correctly classified with fifteen (15) instances wrongly classified and the

positive predicted value is high and this is primarily driven by selection of features that are most relevant to the predictive model problem.

AOD model before feature selection recorded sixty-nine (69) instances correctly classified with 51 instances wrongly classified and we observed that the negative predicted value was very low. While ANFIS model before feature selection recorded thirty-nine (39) instances that were correctly classified with eighty-one (81) instances wrongly classified and we observed that the negative predicted values are low, thereby creating an unnecessary alert of false positive prediction.

VI. CONCLUSION

Chronic kidney disease is the slow loss of kidney function and the foremost form of kidney disease. It normally happens over a period of months to a longtime. Chronic kidney disease is considered a "silent killer" since they are few physical manifestations at its early stages. This study improved on ANFIS model by presenting AOD model for diagnosis of CKD. AOD model employed genetic algorithm for feature selection while hybrid learning algorithm which consists of least square error and conjugate gradient decent for the training.

AOD model was tested with optimized and un-optimized features of the dataset, the same test was carried out on ANFIS model and the result was compared. Finally, the result of the research shown that AOD model outperformed ANFIS model both in convergence and optimum performance. The AOD model converged at the 12th epoch with a minimum error of 0.7333 and a sensitivity of 100%, specificity of 89%, accuracy of 95%, type I error of 6, type II error of 0, type I error rate of 0 and type II error rate of 0.0857 while the ANFIS model converged at the 72nd epoch with minimum error of 0.7333 and a sensitivity of 100%, specificity of 77 %, accuracy of 88%, type I error of 15, type II error 0, type I error rate of 0 and type II error rate of 0.2143. This research shows that AOD model resulted in better classification accuracy and also takes lesser time to converge than ANFIS model. This agrees with the fact that modification of learning algorithm to achieve a minimum training error in few epochs is usually an advantage to improve the performance of classification model. And also can significantly improve comprehensibility of classification problem [31], [11].

REFERENCES

[1] Abhishek, B., Gour, S. M. T., and Gupta, D. (2012). Proposing Efficient Neural Network Training Model for Kidney Stone Diagnosis. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (3), 3900-3904.

[2] AI- Hyari, A., AlTaee, A. M., and Al Taee, M. A. (2013). Clinical Decision Support for diagnosis and management of Chronic Renal Failure, Applied Electrical Engineering and Computing Technologies (AEECT), IEEE Jordan Conference on Pages 1-6.

[3] Akgundogdu, A., Kilic, N., Ucan, O. N., and Akalin, N. (2010). Diagnosis of renal failure disease using adaptive neuro-fuzzy inference system. Journal of Medical Systems, 34: 1003-1009.

[4] Ashraf, M., Chetty, G., and Tran D (2013). Feature Selection Techniques on Thyroid, Hepatitis, and Breast Cancer datasets, International Journal on Data Mining and Intelligent Information Technology Applications (IJMIA), 3(1), 1-8.

[5] Bermingham, M. L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campell, H., Wrigh, A.F., Wilson, J.F., Agakov, F., Navarro, P., and Haley, C.S. (2015). Application of high-dimensional feature selection evaluation for genomic prediction in man.scientific report: A Nature Research Journal, pp 1-12.

[6] Bhatia, S., and Patel, S. (2015). Analysis on different Data mining Techniques and algorithms used in IOTShweta Bhatia Int. Journal of Engineering Research and Applications 5 (1-1):.82-85.

[7] Boukenze, B., Mousannif, H., and Haqiq, A. (2016). Performance of data mining techniques to predict in healthcare case study: chronic kidney failure disease. International Journal of Database Management Systems (IJDMS). 8 (3), 1-9. June.

[8] Cheng-Min, C., Cheng-Tao, Y., and Bor-Wen, C. (2012). Data mining models for diagnosing acute renal failure, The International Journal of Organizational Innovation 4 (4):155-175.

[9] Chowdhury, D. R., Chatterjee, M., and Samanta, R. K. (2011). An Artificial Neural Network Model for Neonatal Disease Diagnosis, International Journal of Artificial Intelligence and Expert Systems (IJAE), 2 (3): 96-109.

[10] Cruz, J., and Wishart, D. (2006). Applications of Machine Learning in Cancer Prediction and Prognosis, Journal of cancer informatics, 2: 59-77.

[11] Enrique, C., Bertha, G., Oscar, F., and Amparo, A. (2006). A Very Fast Learning Method for Neural Networks Based on Sensitivity Analysis, Journal of Machine Learning Research 7, 1159–1182.

[12] Gansevoort, R.T., Correa-Rotter, R., Hemmelgarn, BR., Jafar, TH., Heerspink H.J., Mann, J. F., Matsushita, K., and Wen, C.P. (2013). Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. Lancet 382: 339-352.

[13] Greger, R., Schlatter, E., and Hebert, S.(2001). Milestones in nephrology: Presence of luminal K1, a prerequisite for active NaCl transport in the cortical thick ascending limb of Henle's loop of rabbit kidney. Journal of America Society Nephrology, 12: 1788–1793.

[14] Grzymała-Busse, J., and Hu, M. (2005). A Comparison of Several Approaches to Missing Attribute Values in Data Mining. In: W. Ziarko, Y. Yao, (eds.), Rough Sets and Current Trends in Computing, Lecture Notes in Computer Science, vol., pp. 378–385. Springer Berlin / Heidelberg, 2001.ISBN 978-3-540-43074-2.

[15] Guyon, I., and Elisseeff, A. (2003). An Introduction to Variable and Feature Selection Journal of Machine Learning Research 3:1157-1182.

[16] Hemanth, D.J., Vijila, C.K., and Anitha J. (2010). Application of Neuro-Fuzzy Model for MR Brain Tumor Image Classification, Biomedical Soft Computing and Human Sciences, 16 (1): 95-102

[17] Jang J.R ; Sun C.T; Mizutani E, (1997)."Neuro Fuzzy And Soft Computing". A computational Approach to learning and Machine intelligence. Prentice Hall.

[18] Jebari K., and Madiafi, M. (2013). Selection Methods for Genetic Algorithms.International Journal Emerg Science. IJES, 3(4), 333-344.

[19] Jha, V., Wang, A.Y., Wang, H. (2012). The impact of CKD identification in large countries: the burden of illness. Nephrology Dialysis Transplantation 27: 32-38

[20] Lakany, H., and Conway, B. (2007). Understanding intention of movement from electroencephalograms. Expert Systems, 24(5): 295-304.

[21] Lakshmi. K.R., Nagesh, Y., and VeeraKrishna, M. (2014). Performance Comparison Of Three Data Mining Techniques For Predicting Kidney Dialysis Survivability, International Journal of Advances in Engineering & Technology, Mar., 7(1): 242-254

[22] Luo, C., Zeng, J., Yuan, M., Dai, W., and Yang, Q. (2016).Telco User Activity Level Prediction with Massive Mobile Broadband Data. ACM Transactions on Intelligent Systems and Technology .

[23] Ng, A. Y., and Jordan, M. I. (2001). Convergence rates of the voting gibbs classifier, with application to bayesian feature selection, in 18th International Conference on Machine Learning. Morgan Kaufmann.

[24] Norouzi, J., Yadollahpour, A., Mirbagheri, S.A., Mazdeh, 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, Article ID 6080814, 9 pages

[25] Prechelt, L. (1994). A Study of Experimental Evaluations of Neural Network Learning Algorithms: Current Research Practice, Technical Report 19/94, University of Karlsruhe

[26] Ramya, S., and Radha, N. (2016). Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms, International Journal of Innovative Research in Computer and Communication Engineering, 4(1): 812-820.

[27] Reddenna, L.1., Shaik, S., and Kumar K. (2014). Dialysis Treatment: A Comprehensive Description International Journal of Pharmaceutical Research & Allied Sciences 3(1): 1-13.

[28] Rockville, P., and Suite R. (2009). American Kidney Fund 11921 Perspectives on egfr reporting from the interface between primary and secondary Clinical Journal of America Society of Nephrology; 4(2):258-260.

[29] Rubini, J., and Eswaran, P. (2015). Generating Comparative Analysis of Early Stage Prediction of Chronic Kidney Disease. International Journal of Modern Engineering Research, 5(7): 49-55.

[30] Sun, J., and Reddy, C. (2013).Big data analytics for healthcare.Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining .

[31] Talavera, L, (1999). Feature selection as a preprocessing step for hierarchical clustering. In Proceedings of the 16th International Conference on Machine Learning, ICML-99: 389-397, Bled, Slovenia. I. Bratko and S. Dzeroski, eds. Morgan Kaufmann.

[32] Vijayarani, S., and Dhayanand, S. (2015). Kidney disease prediction using SVM and ANN algorithms International Journal of Computing and Business Research (IJCBR),Volume 6 Issue 2.

[33] Vijayarani, S., and Dhayanand, S. (2015). Data Mining Classification Algorithms For Kidney Disease Prediction"International Journal on Cybernetics & Informatics (IJCI) 4(4): 13-25.

[34] Wagstaff, K. L., and Laidler, V. G. (2005). Making the Most of Missing Values: Object Clustering with Partial Data in Astronomy. Proceedings of Astronomical Data Analysis Software and Systems XIV, 347: 172–176. Pasadena, California, USA.

[35] Zhu, X., Li, X., Zhang, S., Ju, C., and Wu, X. (2017).Robust Joint Graph Sparse Coding for Unsupervised Spectral Feature Selection. IEEE Transactions on Neural Networks and Learning Systems.