Volume 1, Issue IX, September 2022, **No. 19, pp. 249-259**

Submitted 22/9/2022

Final peer reviewed 5/9/2022

Online Publication 5/10/2022

Available Online at http://www.ijortacs.com

A PREDICTIVE ALGORITHM FOR EARLY DETECTION OF NEUROPATHIC DIABETES USING MACHINE LEARNING

[1]Eze Jennifer C., [2]Ezeh Ebere M., [3]Chibueze Kingsley I., [4]Kekong Pius E.

[1]University of Nigeria, Enugu Campus, [2,3]Enugu State University of Science and Technology; [4]Federal University of Health Science, Otukpo.

[1]jeeceze@gmail.com, [2]ebydent@gmail.com, [3]casperine1@gmail.com, [4]piuskekong2019@gmail.com

Abstract

This paper presents a predictive algorithm for the early detection of neuropathic diabetes using machine learning. The study aimed at developing a machine learning model which can classify features of neuropathic diabetes from clinical data and then diagnose the patient. The methodology for the system development is the structural system analysis and structural design model. The machine learning used is the neural network which was trained which data collected from Niger foundation hospital, Enugu, Nigeria, to generate the predictive model which was used to model a neuropathic diabetes prediction system. The result when tested and validated using the tenfold cross validation technique showed an average mean square error performance value of 4.5992e-11 and a regression value of 9.881. The result implied good training performance for the neurons and also good detection of neuropathic diabetes signs. The algorithm when compared with other state-of-the-art predictive algorithms achieved better regression performance with a percentage improvement of 0.08%.

Keywords: Neuropathic Diabetes, Neural Network, Machine Learning, Predictive Model, Diagnose

1. INTRODUCTION

It is no longer news that diabetes is a chronic disease that has continued to affect lives all over the world in a negative way and has remained a major problem as [1] revealed that in the next 15 years about 642 million people will be admitted on diabetes-related cases all over the world. Diabetes is of three major types which are type1, type 2, and gestational diabetes as explained in [2]. Neuropathic diabetes (ND) is the type of diabetes risk factor which is common to the three types of diabetes, with excess glucose in the blood one of the major causes [2]. ND is very devastating as it targets the nerves

[1]Corresponding Author Tel: +234 903 413 5758

connected to the brain and affects them, which in turn affects the operations of other internal organs of the body, especially the motor neuron which is responsible for movement, thus resulting in the amputation of the legs in critical cases, among other risk factors [3]. DN is of four major types which are autonomic, peripheral, mono, and proximity neuropathic respectively; however the peripheral is the most dominant of all diabetes patients [4].

Peripheral diabetes results in the loss of sensors responsible for feelings (i.e the person cannot feel pain), and delay the healing of injuries which often results in the amputation of the affected body parts as the disease keep on increasing [4; 5]. To solve this problem, many solutions have been proposed in [6; 7; 8; 9; 10] using methods such as machine learning and developing a model which can detect the problem, however despite the success, there is a need for a model which can predictive this problem ahead of time for better management. This is because detecting the problem means it has already occurred and hence is not good enough for the diagnosis of a delicate disease like peripheral diabetes, but if the disease can be predicted correctly before it occurs, the patients can be better managed and prevent the risk associated with the problem.

1.2 PROBLEM STATEMENT

Today many works have been developed for the diagnosis of diabetes neuropathic using artificial intelligence techniques, but despite the success, the solutions all focused on detecting the problem after it has occurred which is not very satisfactory and in most cases it is late. This technical problem has resulted in social and economic problems such as amputation, arrhythmia, slow heart rate, vomiting, loss of appetite, urinary retention, eye problems, absence of sweat, hypoglycemia unawareness, and sexual dysfunction, among others. There is a need for a solution that can predict this problem before and alert the patient early enough for fast management and control. The benefit will help reduce the risk implications of the problem of loss of lives.

2. METHODOLOGY

The methodology used for this research is the qualitative analysis which employed open source investigation and simulation method to develop the solution proposed. The investigation allows the researcher to study the impact of peripheral diabetes and identified that early detection is key in managing the problem. The methods involve data collection from a diabetes clinic at Niger foundation, Enugu specialist hospital, neural network, training, and predictive model. The predictive model was then used to predict the problem and system implemented via simulation methodology. The result was evaluated using mean square error and regression and then analyzed.

3. DATA COLLECTION

The source of data collection is the Niger foundation hospital, Enugu, Nigeria which provided data from peripheral neuropathic diabetes samples collected from 120 patients between the age of 15 to 75 and mixed-sex. The geographical location of the hospital is

[1]Corresponding Author Tel: +234 903 413 5758

Latitude 6.443413N and Longitude 7.5144345E. The Google map was presented in figure 3.1. The procedure used for the data collection is fasting lipid profile of the patients, fasting plasma glucose and gycated haemoglobon (HbA1c) and the clinical attributes of the data collected are presented in the table 1;

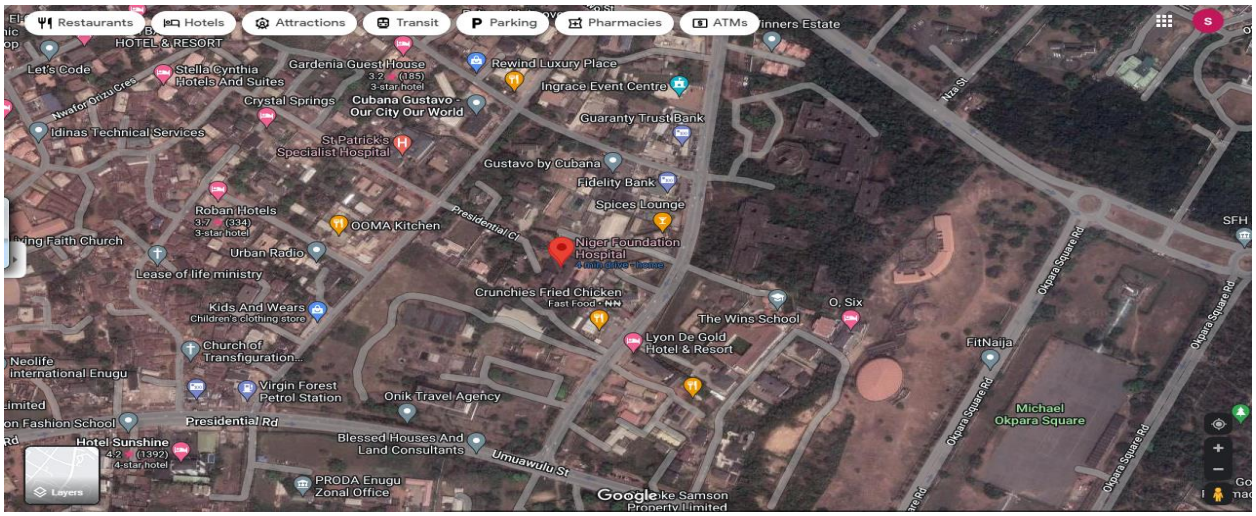

Figure 1: Google location of data source

Table 1: Clinical features of peripheral diabetes

Features	**Values**
Age of patient (yr)	15 to 75
Male	217 (35.62)
Female	395 (64.38)
Diagnosed DPN	218 (35.78)
Course of disease (yr)	8.02 (4.01-13.01)
Body mass index (kg/m²)	24.81 (22.90-27.10)
Waistline (cm)	86.02 (80.01-92.30)
Uric acid (umol/L)	295 (252-340)
Estimated glomerular(mL/min)	58.02 (40.90-78.01)
Systolic blood pressure (mm Hg)	137.02 (124.01-148.01)
Diastolic blood pressure (mm Hg)	77.71+9.69
Fasting blood sugar(mmol/L)	6.81 (5.80-8.33)
Postprandial blood glucose(mmol/L)	11.31 (7.90-14.01)
HbAlc (%)	7.10 (6.31-7.90)
Cholesterol(mmol/L)	4.89+1.02
triglyceride (mmol/L)	1.39 (1.01-1.95)
Low density lipoprotein cholesterol(mmol/L)	1.60 (1.31-1.90)
Low density lipoprotein cholesterol(mmol/L)	1.59 (1.37-1.89)

[1]Corresponding Author Tel: +234 903 413 5758

Blood urea nitrogen (mmol/L)	5.22 (4.39-6.16)

2. DEVELOPMENT OF THE PREDCITVE MODEL

The predictive model was developed using Artificial Neural Network (ANN). The ANN was developed from a simple feed forward Neural Network (FFNN) [11]. The FFNN is made of neuron which has weights, bias and sigmoid activation function. The neurons operation was inspired based on the behavior of neurons in the human brain which when feed with information, can process it and make correct decisions. The FFNN was developed with two hidden layers as shown in the figure 2 which presents the interconnected layers of the neural network with 19 clinical features of peripheral diabetes and the two hidden layers made of 12 neurons in the first layer and 10 neural in the second layer. The neurons are when activated and trained with the data collected was presented in the figure 3;

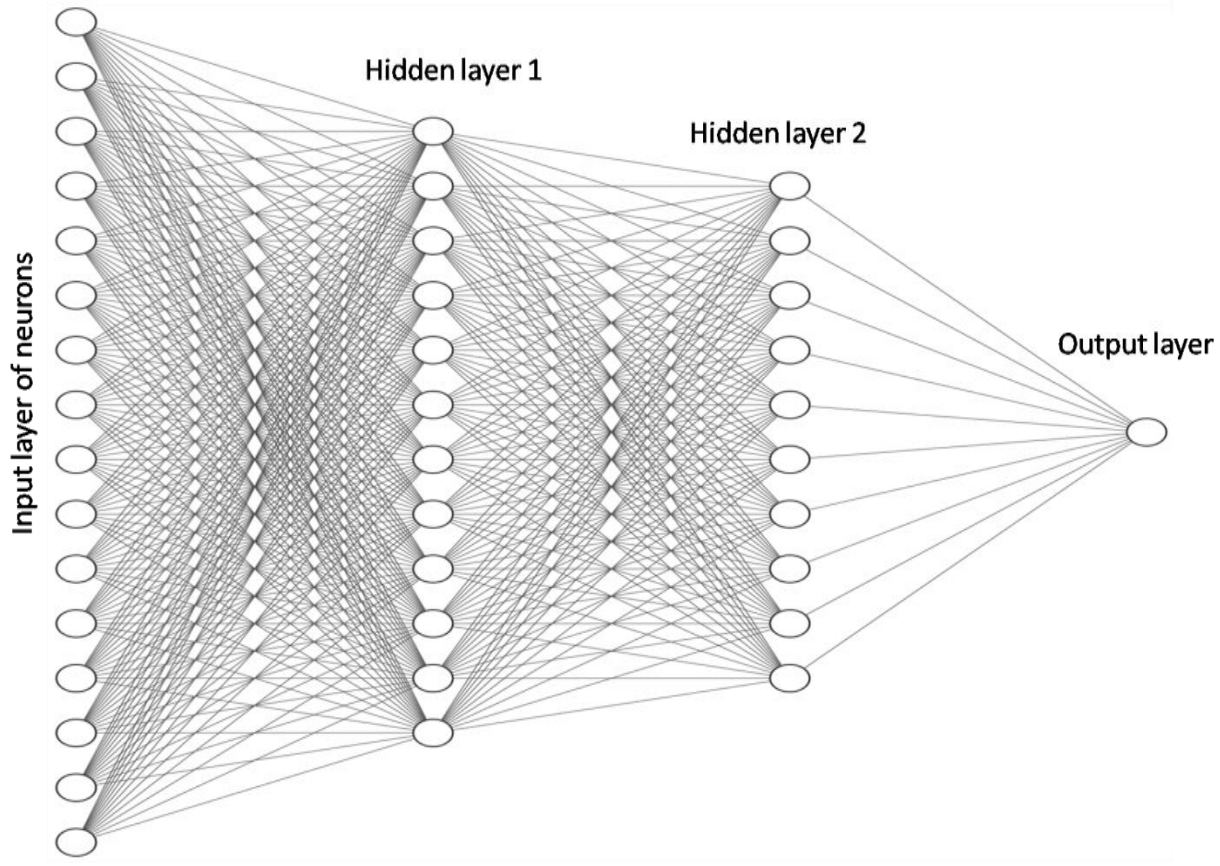

Figure 2: Model of the FFNN

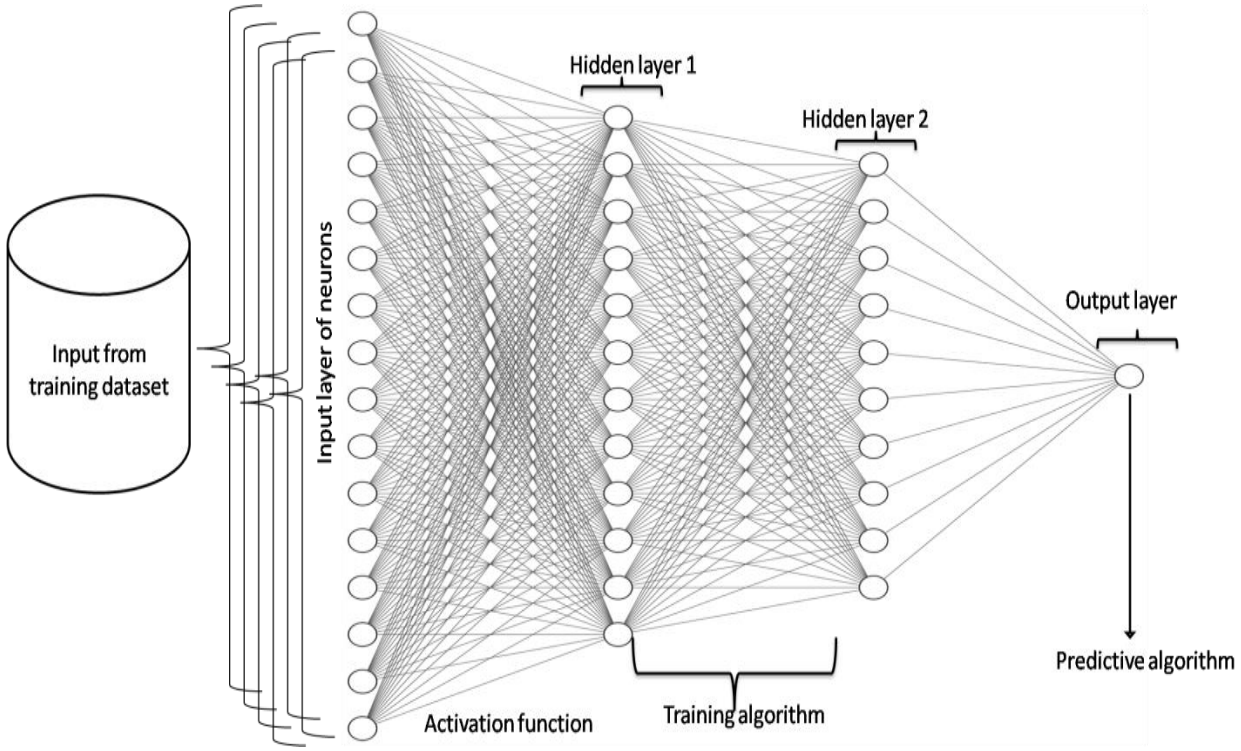

Figure 3: Model of the neural network during training

The figure 3 showed how the neurons was loaded with the clinical features of the peripheral neuropathic diabetes and then activated for training using the back-propagation optimization algorithm adopted from [11]. The training allows the neurons to learn the features via fining the appropriate weights via feedback propagation and then ensuring least loss function, error and then generating the predictive model. Before the training of the FFNN, the data were divided into training, test and validations set by the neural network tool and then train the neurons. The training stops when the neurons have been tested with the features and validated with least error. The pseudocode of the predictive algorithm was presented as;

4.1 PREDICTIVE ALGORITHM

Start
Load data % Clinical features of peripheral neuropathic diabetes
Spit data into 70:20:10 % Training; test and validation sets
Configure FFNN with activation function as in figure 3 % sigmoid function
Activate optimization algorithm % Back propagation algorithm
Set target values % MSE $\simeq$ 0 and Regression (R) $\simeq$ 1
Train FFNN
Is least MSE = true
Check regression
Is optimal R value = true
Check validation % Fivefold validation
Is last five steps relatively constant = true
Stop training % Training ends
Generate Predictive algorithm
Stop

[1]Corresponding Author Tel: +234 903 413 5758

3. IMPLEMENTATION OF THE ALGORITHM IN MATLAB

To implement the predictive model developed to generate a real script which can be used for the development of expert system for clinical diagnosis of peripheral diabetes, the data collected was loaded into a neural network regression application software and then used configure the neural network as shown in the figure 3. The training algorithm was selected from the tool alongside the configuration for the number of layers, before the network was trained. The training was monitored automatically by the system using at epoch until the R and MSE consistently produced fairly constant output values for the loss function and prediction output. The performance of the MSE was used to monitor the error which occurred during the training process until tolerable error was achieved which is the least error tolerable which implied that the training error was minimal and the neurons learn the data due to the impact of the optimization algorithm. The R was also used to test the ability of the algorithm to correctly detect peripheral features from patients clinical data input for test. The MSE was presented in figure 4 and R in figure 5 respectively;

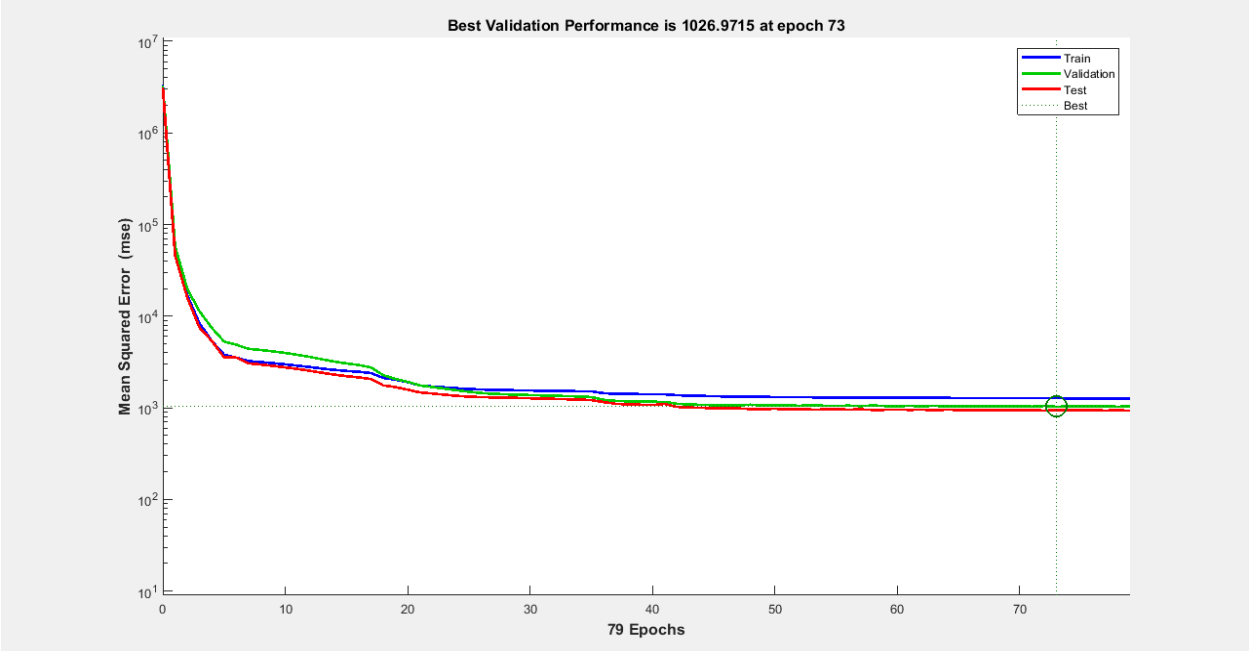

Figure 4: The MSE result of the peripheral predictive algorithm

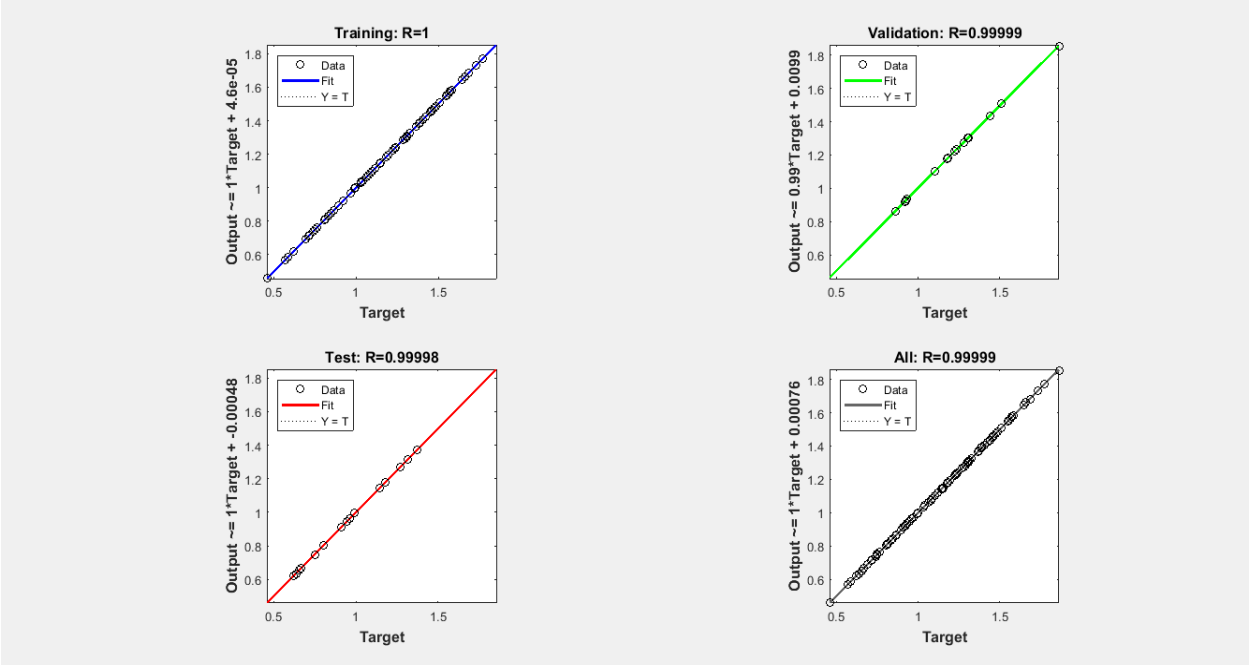

Figure 5: The result of the regression for the peripheral predictive algorithm

4. DISCUSSIONS

The MSE result presented the performance of FFNN during the training process. The aim was to achieve a MSE equal or approximately 0 which implied that the training was perfect without overshot. The result achieved for the MSE is 4.9392e-11 at epoch 73 which implied that the neurons achieved the least error at epoch 73. To evaluate the performance of the predictive model in early detection of peripheral diabetes, the regression result was used to test the capacity of the algorithm to correctly predictive the attributes of peripheral diabetes on patients using the relationship between true positive rate and false positive rate. The aim of the R was to achieve a value equal or approximately 1 which implied correct prediction of peripheral disease. From the result in figure 5, the R achieved overall value of 0.9999 which implied that the predictive algorithm was able to correctly detect the features of the disease early on the patients for diagnosis. The results were validated using tenfold cross validation model in [11] and the results obtained was presented in the table 2;

Table 2: Result of the predictive algorithm Validation

Fold	Epoch	MSE	Regression
1	68	4.6462e-11	0.9999
2	74	3.9342e-11	0.9823
3	54	3.3432e-11	0.9814
4	78	4.7552e-11	0.9852
5	75	3.4345e-11	0.9911

[1]Corresponding Author Tel: +234 903 413 5758

6	56	5.4333e-11	0.9894
7	60	5.0534e-11	0.9767
8	81	4.6632e-11	0.9934
9	74	4.7893e-11	0.9999
10	67	5.8642e-11	0.9912
Average	68.7	4.5992e-11	0.9891

The table 2 presented the validation performance of the algorithm developed for the early detection of peripheral diabetes. The result achieved an average of 4.5992e-11 as MSE and regression of 0.9891 at an average epoch of 68.7. The result implied that the predictive algorithm will correctly detect peripheral features of neuropathic diabetes in patients. The algorithm was compared with other state of the art machine learning based algorithms and the result of the regression was presented in table 3;

Table 3: Comparative prediction algorithm regression

Techniques	Regression
CNN [13]	0.90
XG boast [14]	0.89
FFNN [14]	0.86
Logistic Regression [14]	0.83
LASSO [14]	0.83
New FFNN	0.98

The table 3 presented the comparative results obtained for regression performance of machine learning based predictive algorithms. The results were analyzed using the graph in figure 5;

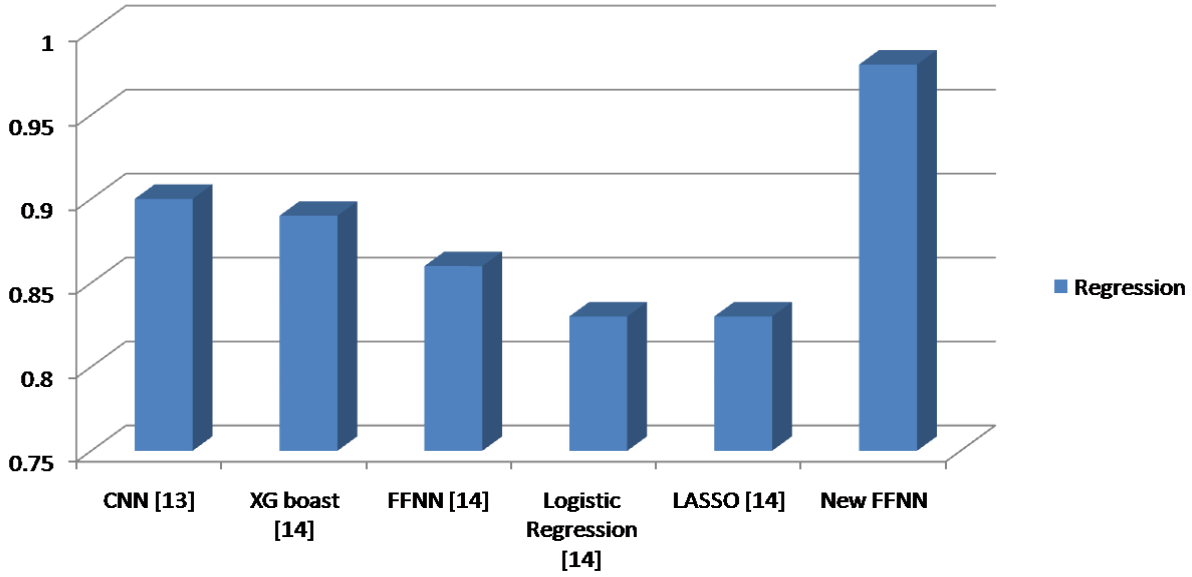

Figure 5: Comparative analysis

The figure 5 presented comparative regression performance of the prediction algorithm developed with other stat of the art machine learning algorithms. The result showed that the new FNN was able to achieve highest regression performance when compared with the counterparts with a percentage improvement of 0.08%. This was due to the deep configuration of the neurons to form deep layers which ensured multiple computation process using the optimization based back propagation algorithm.

5. Conclusion

Peripheral neuropathic diabetes has remained a major challenge over the years all around the world due to its risk factors which has handicapped countless number of persons. The research have developed and provided an algorithm which can be used to develop an expert system for early detection of the disease. The algorithm was tested and validated. The result showed that it can correctly predict peripheral diabetes from patient clinical samples. The output of the result will allow fast diagnosis of the patients and early management to prevent implication and risk factors identified in the problem statement.

6.1 Contribution to knowledge

- The paper presented a neural network based predictive algorithm for early detection o peripheral diabetes.

6.2 Recommendation

The paper provided open source algorithm which can be used for the development of expert system for the early detection and diagnosis of peripheral diabetes.

6.3 Ethical Consideration

The author declares no conflict of interest.

REFERENCE

1. Zou, Q., Qu, K., Luo, Y., Yin, D., Ju, Y., & Tang, H. (2018). Predicting Diabetes Mellitus with Machine Learning Techniques. *Frontiers in Genetics*, 9. doi:10.3389/fgene.2018.00515
2. Meijer J, Lange F, Links P, van der Hoeven H. (2008) “Muscle fiber conduction abnormalities in early diabetic polyneuropathy”; Clin Neurophysiol; 119: 1379–1384.
3. Lyssenko, V., Jonsson, A., Almgren, P., Pulizzi, N., Isomaa, B., Tuomi, T., . . . Groop, L. (2008). Clinical Risk Factors, DNA Variants, and The Development of Type 2 Diabetes. The New England Journal of Medicine, 359, 2220-2232. doi:10.1056/NEJMoa0801869
4. Lu B, Hu J, Wen J, Zhang Z, Zhou L, Li Y, et al. Determination of peripheral neuropathy prevalence and associated factors in Chinese subjects with diabetes and pre-diabetes:
5. Yang CP, Lin CC, Li CI, Liu CS, Lin WY, Hwang KL, et al. Cardiovascular risk factors increase the risks of diabetic peripheral neuropathy in patients with type 2 diabetes mellitus: the Taiwan Diabetes Study. Medicine (Baltimore) 2015;94:e1783.
6. Tesfaye S., Boulton J., Dyck P., Freeman, M. Horowitz, P. Kempler, G. Lauria, R. A. Malik, V. Spallone, Vinik

A. (2010) “Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments,” Diabetes care, vol. 33, no. 10, pp. 2285–2293, 2010.

7. Tetiana Dudkina, IevgenMeniailov, KseniiaBazilevych, Serhii Krivtsov and Anton Tkachenko (2021). Classification and Prediction of Diabetes Disease using Decision Tree Method. Creative Commons License Attribution 4.0 International (CC BY 4.0). CEUR Workshop Proceedings (CEUR-WS.org)
8. Sonar, P., &JayaMalini, K. (2019). Diabetes prediction using different machine learning approaches. *In 2019 3rd International Conference on Computing Methodologies and Communication* (ICCMC), 367-371.
9. Sowjanya, K., Singhal, S. and Choudhary, C. "MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices." pp. 397-402.
10. Sarki, R., Michalska, S., Ahmed, K., Wang, H., & Zhang, Y. (2019). Convolutional neural networks for mild diabetic retinopathy detection: an experimental study. *bioRxiv*, 763136.
11. Mba J.C. and Asogwa T.C (2022)” Modeling of neuro based strategy for mitigating cyber threat in 4G wireless network using artificial intelligence technique” [J] IJAICCSS; vol 1; issue 7; pp. 2-16.
12. Saxena, K. and Sharma, K. "Diabetes mellitus prediction system evaluation using C4. 5 rules and partial tree." pp. 1-6.
13. Jin, Y., Li, F., Vimalananda, V. G., & Yu, H. (2019). Automatic detection of hypoglycemic events from the electronic health record notes of diabetes patients: empirical study. *JMIR medical informatics*, 7(4), e14340.
14. Li, L., Lee, C. C., Zhou, F. L., Molony, C., Doder, Z., Zalmover, E., ... & Wu, C. (2021). Performance assessment of different machine learning approaches in predicting diabetic ketoacidosis in adults with type 1 diabetes using electronic health records data. *Pharmacoepidemiology and drug safety*, 30(5), 610-618.