International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 6 Issue 7, November-December 2022 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Development of a Hybrid Dynamic Expert System for the Diagnosis of Peripheral Diabetes and Remedies using a Rule-Based Machine Learning Technique

Omeye Emmanuel C.[1], Ngene John N.[2], Dr. Anyaragbu Hope U.[3],
Dr. Ozioko Ekene[4], Dr. Iloka Bethram C.[5], Prof. Inyiama Hycent C.[6]

[1, 2, 3, 4, 5, 6]Department of Computer Science,
[1, 3, 5]Tansian University, Umunya, Anambra State, Nigeria
[2, 4, 6]Enugu State University of Science and Technology, Enugu, Nigeria

ABSTRACT

This paper presents the development of a hybrid dynamic expert system for the diagnosis of peripheral diabetes and remedies using a rule-based machine learning technique. The aim was to develop a solution to the risk factors of peripheral diabetes. The methodology applied in this study is the experimental method, and the software design methodology used was the agile methodology. Data was collected from Nnamdi Azikiwe University Teaching Hospitals (NAUTH) and the Lagos State University Teaching Hospital (LASUTH) for patients between the ages of 28-87years suffering from peripheral neuropathy. Other methods used were data integration by applying uniform data access (UDA) technique, data processing using Infinite Impulse Response Filter (IIRF), data extraction with a computerized approach, machine learning algorithm with Dynamic Feed Forward Neural Network (DFNN), rule-base algorithm. The modeling of the hybrid dynamic expert system and remedies was achieved using the DFNN for the detection of DPN and a rule-based model for remedies and recommendations. The models were implemented with MATLAB and Java programming languages. The result when evaluated achieved a Mean Square Error (MSE) of 4.9392e-11 and Regression (R) of 0.99823. The implication of the result showed that the peripheral diabetes detection model correctly learns the peripheral diabetes attributes and was also able to correctly detect peripheral diabetes in patients. The model when compared with other sophisticated models also showed that it achieved a better regression score. The reason was due to the appropriate steps used in the data preparation such as integration and the use of IIFR filter, feature extraction, and the deep configuration of the regression model.

How to cite this paper: Omeye Emmanuel C. | Ngene John N. | Dr. Anyaragbu Hope U. | Dr. Ozioko Ekene | Dr. Iloka Bethram C. | Prof. Inyiama Hycent C. "Development of a Hybrid Dynamic Expert System for the Diagnosis of Peripheral Diabetes and Remedies using a Rule-Based Machine Learning Technique" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-6 | Issue-7, December 2022, pp.429-438, URL: www.ijtsrd.com/papers/ijtsrd52356.pdf

IJTSRD52356

KEYWORDS: *Expert System; Peripheral Diabetes; Neural Network; Machine Learning, MSE, R*

1. INTRODUCTION

The International Diabetes Federation (IDF) suggested that diabetes mellitus has reached the proportion of epidermis worldwide due to its high rate of growth from 425 million people worldwide in 2017 and is predicted to get to 628 million by 2045 (International Diabetes Federation, 2017). As this number rises, the prevalence and complications of the disease rise too. The common cause of neuropathy worldwide is Diabetic Peripheral Neuropathy (DPN) and it has been estimated to affect about half of the people suffering from diabetes worldwide (Thibault et al., 2016).

DPN has been identified to be the primary cause of the impairments of quality of life, morbidity, and increased mortality. It affects the distal foot, and toes and slowly goes on to affect the leg and feet proximally to involve the legs in a stocking distribution. It can also be known to lead to losses in nerve fibers which goes on to affect the somatic and autonomic divisions, which then cause the occurrence of diabetic retinopathy and nephropathy. One of the main clinical consequences of DPN is foot ulceration and painful neuropathy (Albers and Pop-Busui, 2014; Alleman et al., 2015).

Diabetic Neuropathic Pain (DNP) is characterized by sensations of burning, shooting, sharp and lancinating, or even electric shock and tingling. And all these uncomfortable sensations are usually worse at night which later causes disturbance to sleep. These pains can be constant which subsequently affects the quality of life negatively (Akter, 2019). Therefore, there is a need for serious measures to be taken to mitigate the development and dominance/prevalence of this disease. Managing diabetes requires the monitoring and control of insulin levels in the body to prevent this nerve damage and ensure slow progress, however, prevention is better than caused which presents the need for early prediction (Moaz et al., 2011).

Machine Learning (ML) is an artificial intelligence technique that can learn and make correct decisions. This ML consists of various regression algorithms which can be trained to make predictions and have been used over time to diagnose diabetes (Zhang et al., 2019; Li et al., 2021), however, solutions have not been obtained for DPN and this will be achieved in this study using Artificial Neural Network (ANN). ANN is a set of neurons that can learn and solve time series problems, this will be used to develop a regression algorithm that will be able to predict DN on patients for remedies.

2. LITERATURE REVIEWS

Jian et al., (2021) presented a machine-learning approach to predicting diabetes complications. The study applied several supervised classification algorithms for developing different models to predict and classify eight diabetes complications such as metabolic syndrome, obesity, dyslipidemia, neuropathy, diabetic foot, nephropathy, hypertension, and retinopathy. The datasets used for the study were collected from the Rashid center for diabetes. The algorithms applied in the study include SVM, AdaBoost, XGBoost, random forest, decision learning, and logistic regression. From the results presented, it is observed that random forest, AdaBoost, and XGBoost have better performance accuracy than others.

Shahabeddin et al., (2019) presented a study on artificial intelligence applications in type 2 diabetes mellitus care. It is aimed to identify the artificial intelligence (AI) application that can be effective for the care of type 2 diabetes mellitus. The variables used were BMI, fasting blood sugar, blood pressure, triglycerides, low-density lipoprotein, high-density lipoprotein, and demographic variables, while the algorithms commonly used are SVM and naïve Bayes. The result shows that SVM algorithms had higher performance accuracy than the others.

Reshma and Anjana (2020) presented a study on a literature survey on different techniques used for predicting diabetes mellitus. The research reviews the application of neural networks/multilayer perceptron, logistic regression, AdaBoost, random forest, Instance-Based-K (IBK), J48, SVM, Naïve Bayes, and random forest for the prediction of diabetes mellitus.

Abdullah et al. (2011) presented a study on an expert system for determining diabetes treatment based on cloud computing platforms. The research surveyed some of the state-of-the-art methods applied for the purpose of determining the treatment of diabetes and then developed a cloud-based expert system with Google App Engine. The solution was effective but very complex to improve on.

3. METHODOLOGY

The methodology applied for the development of this work is the experimental method, the software development life cycle used is the agile model while the software design methodology used is the adopted Agile methodology. This methodology was aimed at the prompt delivery and aligning project goals to the business needs. It features four iterative phases of feasibility and business study, functional and mathematical models, implementation, and simulations. The agile model also features detailed documentation which is lacking in most other frameworks. It involves an iterative and incremental approach that emphasizes continuous user involvement with timely delivery of the new system within the specified work plan, time, and budget. The methods are data collection, data integration, data processing, data extraction, machine learning algorithm, rule base algorithm, and hybrid dynamic expert system.

Data Collection and Procedure

The primary source of data collection is the NAUTH, which provided data on DPN while the secondary source of data collection is the Lagos State University Teaching Hospital (LASUTH) which provided data on peripheral diabetes. The sample size of data collected from NAUTH was provided by 524 patients between the ages of 28 to 87yrs of whom 46.6% are male and 53.4% are female.

The procedure for data collection was executed by Oguejiorfor et al. (2019) considering various tests conducted on the peripheral patients using the fasting lipid profile of the patients, fasting plasma glucose, and glycated hemoglobin (HbA1c) respectively to collect data for three months which was collected and used for this study as the primary data. The data collection considered key attributes for peripheral neuropathy like age, duration of diabetes mellitus, body mass index, waist circumference, and hemoglobin percentage.

The LASUTH provided data of peripheral diabetes from 225 patients between the ages of 28 to 87yrs collected over a period of 3 months. The procedure for data collection used biothesiomtry for the collection of peripheral arterial disease data is ankle brachial pressure induced as posited in Anthonia et al. (2015). The data collection considered general DPN attributes such as age, elevated level of hemoglobin, gender, duration of diabetes, body weight, history of hypertension, low density of lipoprotein cholesterol, triglyceride, and total cholesterol.

Data Integration

Data integration is a process of merging data collected from various sources and was performed in this research. The data collected from the primary and secondary sources of data collection were integrated together to develop the training dataset required for this work. The approach used for the data integration is the using Uniform Data Access Integration (UDAI) technique (Abe, 2021) which ensures that data are arranged in a consistent format for ease of extraction.

Data Processing

The quality of training data is vital for the reliability of the algorithm which will be developed with it. To achieve this, the data integrated was processed to remove noised from the data attribute in the data initiated from the data collection instrument using finite and Infinite Impulse Response Filter (IIRF) developed by Wayne (2010) for the continuous processing of diabetes data adopted and used to process the data collected.

Machine Learning Model

The ML algorithm adopted for the development of the regression model is the Dynamic Feed-Forward Neural Network (DFNN). The reason this algorithm was used ahead of other ML algorithms was due to its ability to precisely identify system input as an autoregressive model and then train to develop the regression algorithm. This DFNN was developed using interconnected neurons inspired by the attributes of the training set, tanh activation function, and back-propagation which was used to train the neurons until the DPN data were learned and then generate the regression algorithm used for step-ahead prediction of DPN.

Rule-base Model

The rule-based model was developed using an optimization approach that employed logical conditions to develop the remedies for the output of the DPN predicted. This rule-based model was developed using data generated from a series of consultations conducted with domain experts from the primary and secondary source of data collection, and also from international standards established by the World Health Organization (WHO) were used to develop remedies that work in collaboration with the result of the prediction model.

The Hybrid Model of the Dynamic Expert System

The hybrid model of the dynamic expert system was developed using the DFNN model developed for the prediction of DPN and also the rule-based model developed for the remedies. The two models were married as hybrid and then used to develop the dynamic expert system. The dynamic nature of the expert system was inherited from not only its high-level intelligence but also its ability to diagnose both peripheral neuropathic diabetes and also autonomic neuropathic diabetes and make recommendations that will help the patients manage the problem at a very early stage to prevent other severe nervous implications. The system with few inputs from the patients was able to predict DPN and then offer solutions in form of remedies which will allow the patients to manage the problem and leave a better life

4. THE SYSTEM DESIGN

The processed with the modeling of the data collected. The data collected attributes for diabetic peripheral neuropathy (DPN) as shown in the data description table 1;

Table 1: Clinical Data for Peripheral diabetes

Features	Values
Age (yr)	25.02-87.02
Male	217 (35.61)
Female	395 (64.39)
Diastolic blood pressure (mm Hg)	76.70+9.09

Fasting blood sugar(mmol/L)	6.70 (5.70-8.23)
Postprandial blood glucose(mmol/L)	10.30 (7.90-13.91)
Diagnosed DPN	218 (34.38)
Course of disease (yr)	8.01 (4.01-13.01)
Body mass index (kg/m²)	25.10 (23.10-28.90)
Waistline (cm)	86.01 (80.11-91.90)
Systolic blood pressure (mm Hg)	136.01 (123.01-147.01)
HbAlc (%)	6.90 (6.01-7.02)
Cholesterol(mmol/L)	4.79+1.03
triglyceride (mmol/L)	1.29 (1.01-1.94)
Low density lipoprotein cholesterol(mmol/L)	1.50 (1.22-1.80)
Low density lipoprotein cholesterol(mmol/L)	1.49 (1.28-1.79)
Blood urea nitrogen (mmol/L)	5.11 (4.29-6.06)
Uric acid (umol/L)	293 (251-339)
Estimated glomerular(mL/min)	57.01 (39.90-77.01)

Development of the Machine Learning Regression Model using Dynamic Neural Network

The regression model for the prediction of the DPN was developed using a DFNN. The DFNN was developed from a simple feed-forward neural network via the addition of feedback at the output layer to the input layer. The reason was to improve the training performance and learning accuracy via the short-term memory of the neurons. The configuration of the FFNN was adopted from Aishwarya and Vaidehi (2019) and reconfigured to develop the new regression model. The FFNN has weight, bias, activation function, and training algorithm as shown in figure 1;

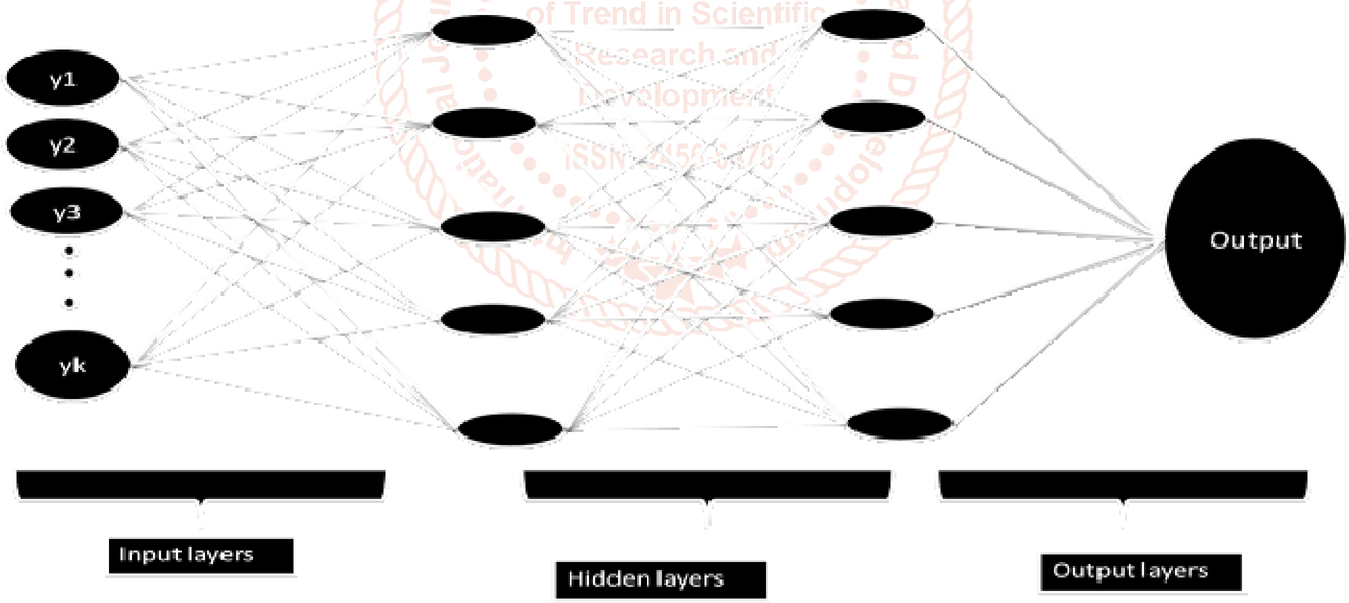

Figure 1: Architectural Diagram of the FFNN

Figure 1 presented the structure of the FFNN model with the interconnected neurons (yk). The neurons each are activated with the tansig activation function which ensured that each neuron triggers and give output within the range of -1 and +1. The neurons are interconnected in two multi-layered for to form the hidden layers and then the final output where the activation vector is identified for training. However, in this case, the FFNN was remodeled to DFNN using a feedback approach achieved with a back-propagation algorithm and the parameters obtained from the training set which is 32 input neurons with each representing features of DPN as shown in figure 2;

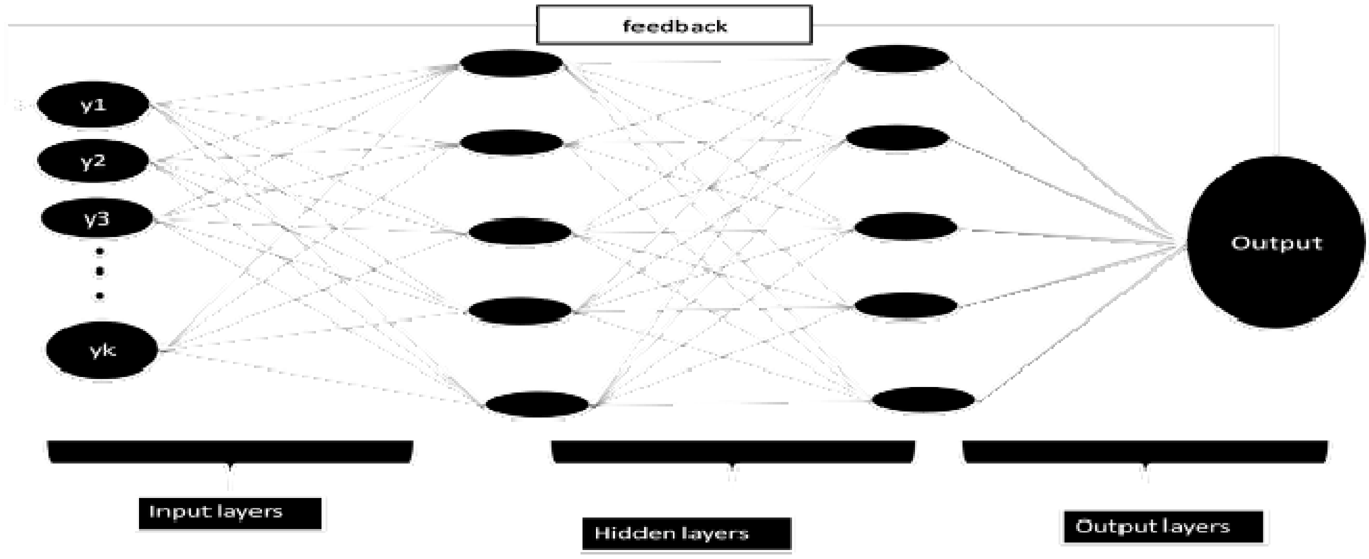

Figure 2: The Architectural Model of the DFNN

Figure 2 presented the model of the DFNN which was developed by the backpropagation algorithm adopted from Mekaoui et al. (2013) to allow the output feedback to the input neurons until the data is learned. To train the neuron, the data collected was loaded into the DFNN architecture and then trained using the gradient descent training algorithm (Staudemeyer and Morris, 2019)

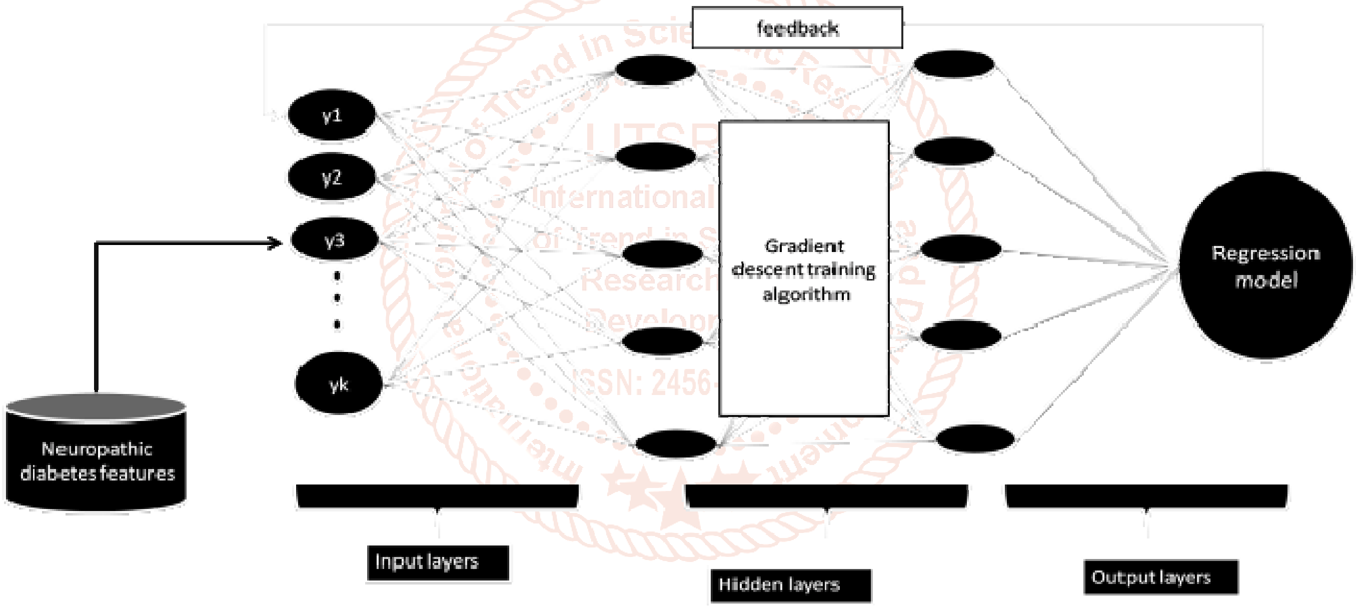

Figure 3: The DN Regression Model Developed

Figure 3 presented the DPN regression model developed which was used for the prediction of DPN in patients. The flow chart of the regression model was presented in figure 4;

Figure 4 presented how the regression model for the prediction of DPN was developed. The data collected were processed using the IIRF and then extracted by the computerized feature extraction approach into the DFNN model develop with a back propagation algorithm for training, during the training process, the neurons were adjusted by the gradient descent to learn the features of the DPN until the Least Mean Square Error (MSE) was achieved, which implied good training performance and then generated the regression model.

Development of a Rule-based Model using for Remedies and Recommendations

This section discusses the remedies for the output of the regression model. If the model output implied that the patient has signs of DPN, then certain recommendations inspired by domain experts in the treatment of DPN were used to develop the rule-based for remedies and recommendation

Development of a Dynamic Expert System Model

Having developed the regression model for the detection of DPN and also the rule-based model for the recommendation, the two models were used to develop the hybrid model of the dynamic expert system as shown in figure 5;

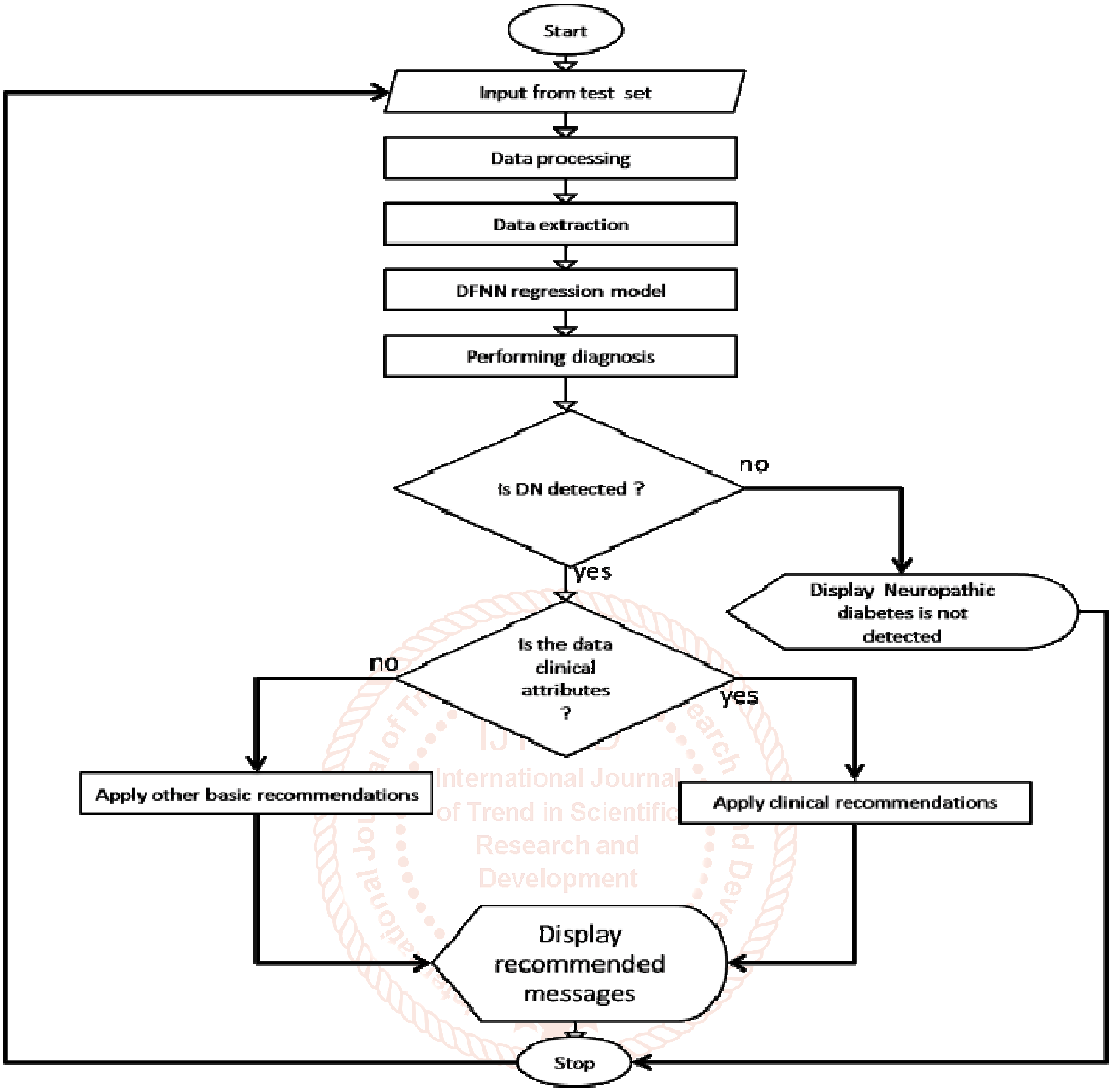

Figure 5: Flowchart of the Dynamic Expert System

From the flowchart of figure 5, the input test data from patient health records were loaded into the system and the features were extracted based on the computerized feature extraction techniques and then feed forward to the DFNN regression model developed to diagnose the patients by checking for both clinical and basic features of DPN. If the output is positive, then recommendations are made based on the nature of the DPN attributes detected.

5. SYSTEM IMPLEMENTATION

The system developed was implemented using Matlab and java programming language. In the Matlab environment, neural network application programming software was called and then used to upload the data collected from hospitals for configuration and training until the best training result was achieved, and then the DPN regression model was generated. The generated algorithm was exported to the java programming platform (Netbean) and then used to develop the dynamic expert system as shown in figure 6;

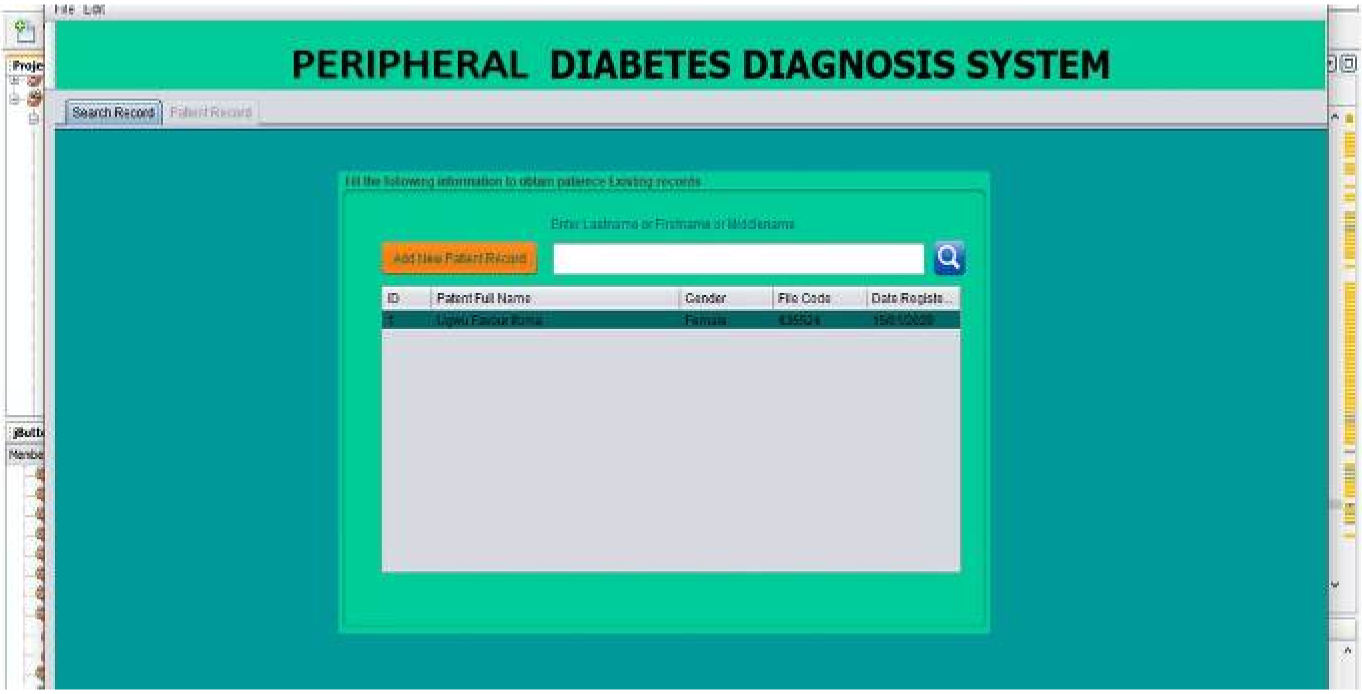

Figure 6: Interface Result of the Expert System for Diagnosis of Diabetes

Figure 6 presented the interface result of the expert system developed for the diagnosis of DPN using the data generated from the dynamic neural network training process.

6. RESULTS

The section presented the result of the training process for the DFNN and then the result of the expert system application for the diagnosis of diabetes. Figure 7 presented the result of the training process using Mean Square Error (MSE) and Regression (R).

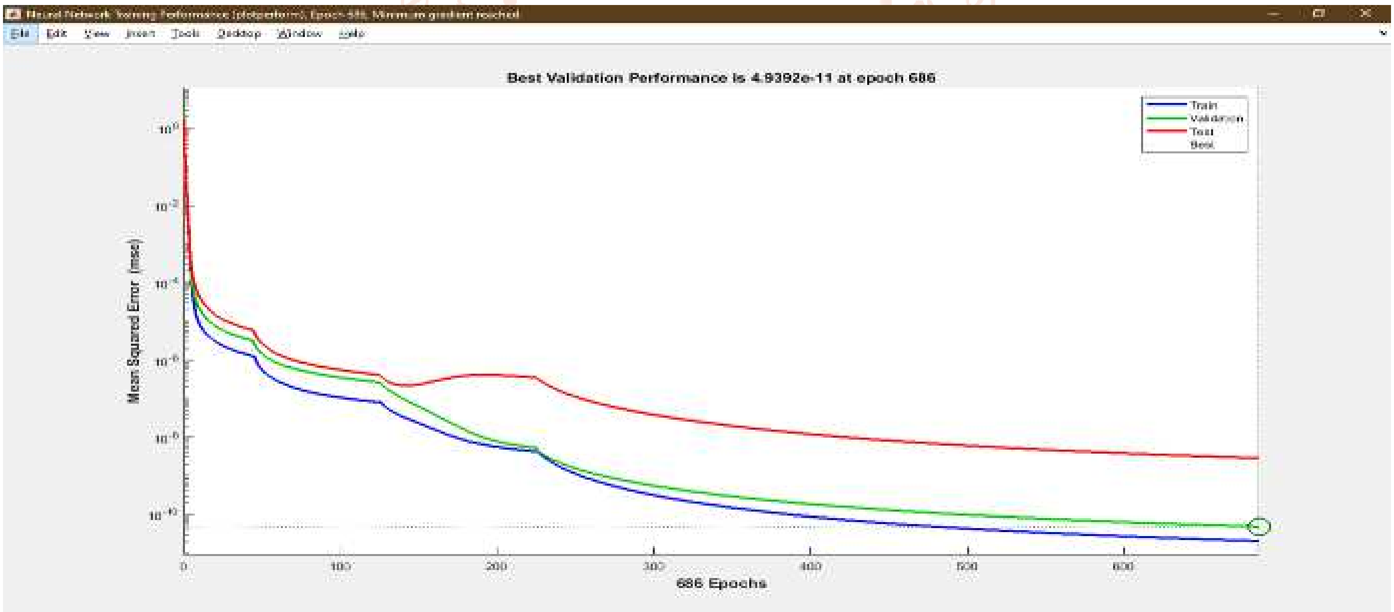

Figure 7: Result of MSE Result

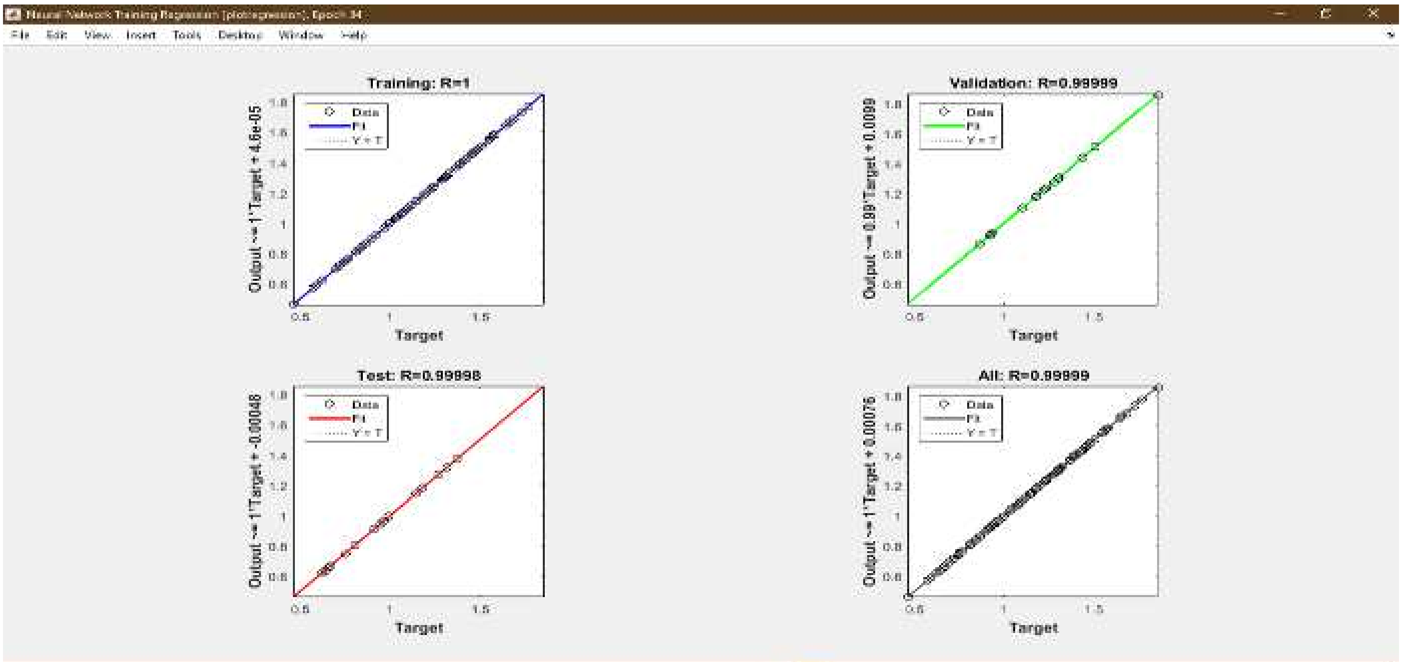

Figure 8: Regression Result

Figure 7 presented the MSE result of the training process. The result showed that MSE of 4.9392e-11 and at epoch 686. The Regression result in figure 8 is 0.99823. The MSE result implied that the training error was minimal and tolerable as it is approximately zero. The regression result also showed that the model generated was able to predict DPN correctly. Having evaluated the algorithm and achieved a good result for the MSE and R, it was used to develop the expert system and tested in figure 9;

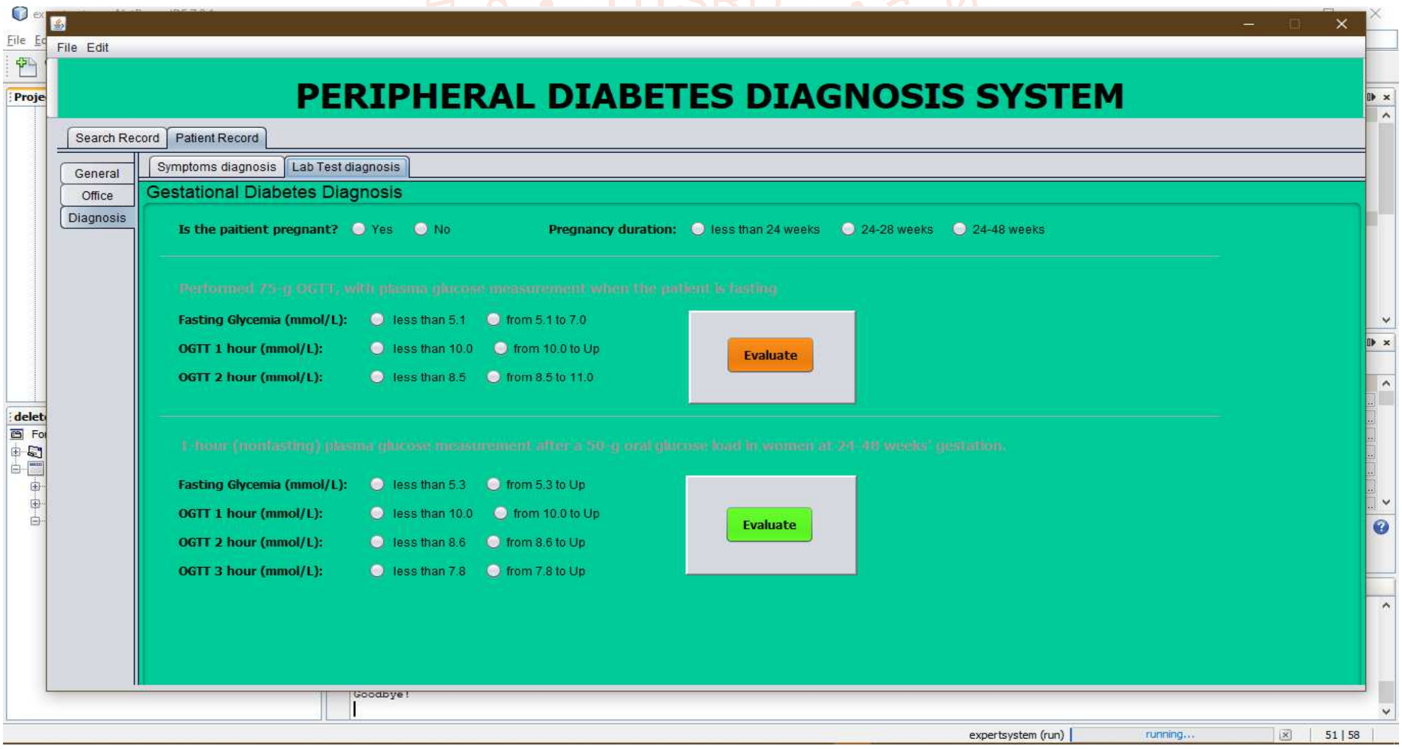

Figure 9: Result of the Interface for Clinical Attributes Diagnosis

Figure 9 presented the performance of the clinical diagnosis interface developed with the regression model generated. The result presented the input parameter which considers basic DPN attributes for the diagnosis of patients. The result when applied for the diagnosis of DPN was presented in figure 10;

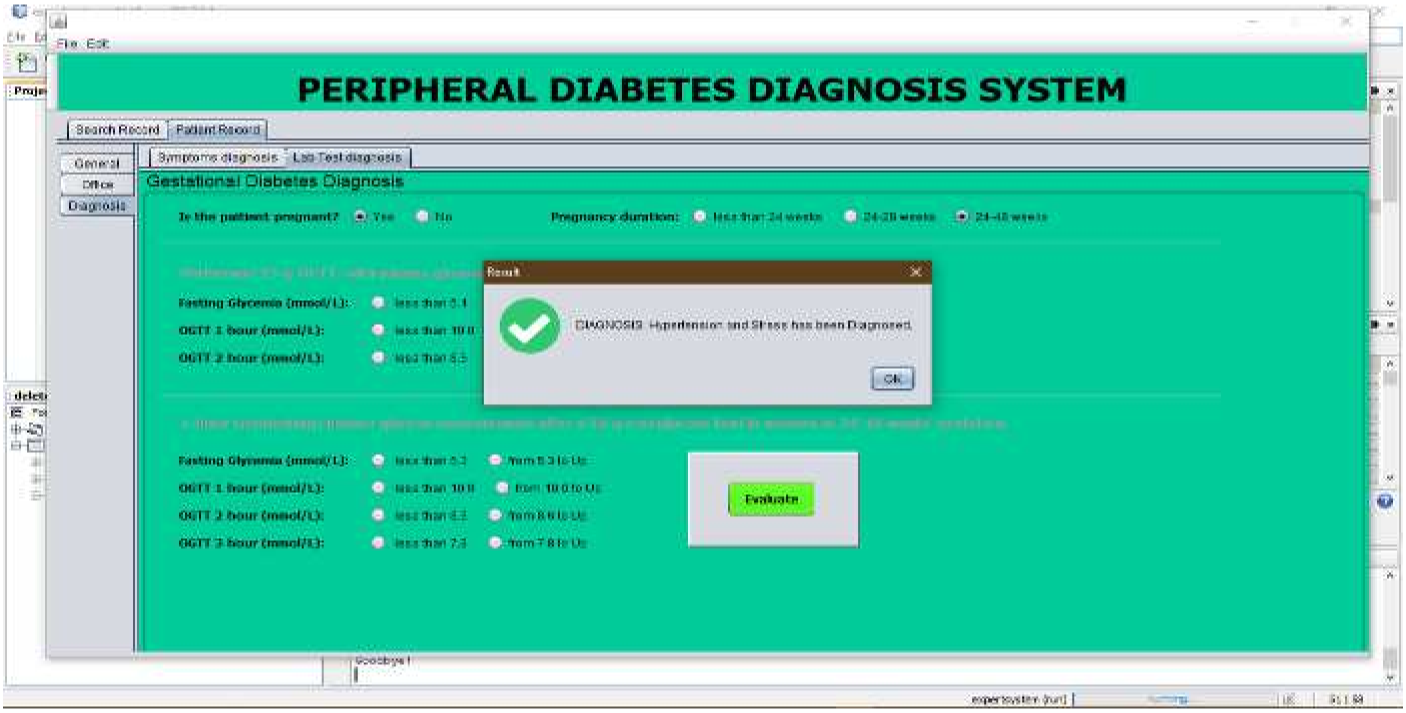

Figure 10: Result of the Diagnostic Process with DN Attributes

Figure 10 showed how the diagnostic system was used to train and detect DPN based on the clinical attributes of the patient. To achieve this, the regression model identifies the inputs from the patient and then trains the data to make the decision on the DN status of the patient. The output of the result was used by the rule-based model for recommendation as shown in the results of figure 11;

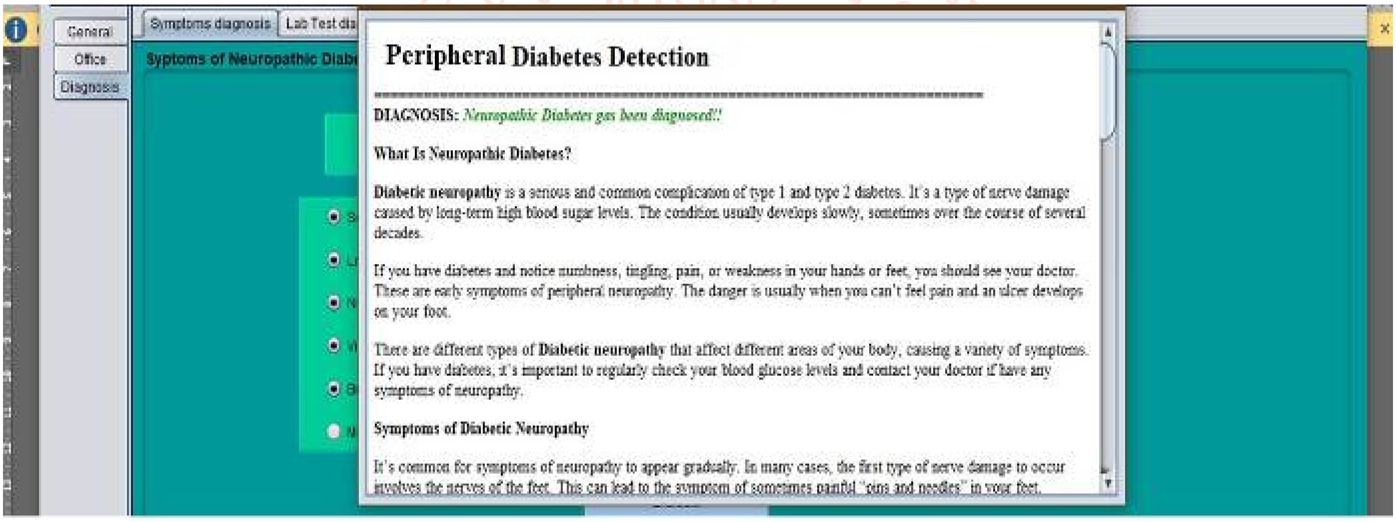

Figure 11: Results of the Recommendation

Figure 11 presented the recommendation performance based on the rule-based model to give remedies to patients and help the patient manage the student.

7. CONCLUSION

This work presented a dynamic system for the diagnosis of DPN. This was achieved after the literature review identified peripheral diabetes as a major cause of amputation among patients. The review also uncovered that solutions have not been obtained for a dynamic system to diagnose DPN. The method used for the new system development is data collection, data integration, data processing, data extraction, machine learning algorithm, rule base algorithm, and hybrid dynamic expert system. The modeling of the DPN detection system was achieved using DFNN and data collection to develop the regression model which was then used to model the dynamic diagnostic system. The system was implemented with Matlab and Java programming languages. The result when evaluated achieved an MSE of 4.9392e-11 and R of 0.99823. The implication of the result showed that the regression model correctly learns the DPN attributes and was also able to correctly detect DPN from patients. The regression model when compared with other sophisticated models also showed that it achieved a better regression score. The reason was due to the appropriate steps used in the data preparation such as integration and the use of IIFR filter, feature extraction, and the deep configuration of the regression model. All these ensured good training performances and also a good regression model which was able to correctly detect DPN.

REFERENCES

[1] Abdullah A., Majda A., Fatimah M., Najwa K., (2011). "An Expert System of Determining Diabetes Treatment Based on Cloud Computing Platforms". (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5), 2011, 1982-1987

[2] Abe D (2021) "Five types of data integration" Article; available at http://www.integrate.io. Access6/29/2022.

[3] Akter N., (2019). "Diabetic Peripheral Neuropathy: Epidemiology, Physiopathology, Diagnosis and Treatment", Assistant Professor (Endocrinology & Metabolism), Dept. of Medicine, MARKS Medical College & Hospital, Mirpur, Dhaka, Bangladesh. Pp 35-48

[4] Albers J, Pop-Busui R. (2014). "Diabetic Neuropathy: Mechanisms, Emerging Treatments, and Subtypes". Curr Neurol Neurosci Rep. 2014;14:473

[5] Alleman C., Westerhout K., Hensen M., Chambers C., Stoker M., Long S, (2015). "Humanistic and Economic Burden of Painful Diabetic Peripheral Neuropathy in Europe: a Review of the Literature. Diabetes", Res Clin Pract. 2015; 109(2):215-25.

[6] Anthonia O., Olufunmilayo A., Babatunde S.,& Alfred A., (2015) "Screening for peripheral neuropathy and peripheral arterial disease in persons with diabetes mellitus in a Nigerian University Teaching Hospital" Res Notes (2015) 8:533 DOI 10.1186/s13104-015-1423-2

[7] International Diabetes Federation. (2017) "IDF Diabetes Atlas. 8th ed.", Brussels, Belgium: International Diabetes Federation; 2017

[8] Jian Y., Pasquier M.,Sagahyroon A.,Aloul F. (2021). "A Machine Learning Approach to Predicting Diabetes Complications". Healthcare, 1712. https://doi.org/10.3390/healthcare9121712

[9] Li L., Lee C., Zhou F., 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.

[10] Moaz A., Moutasem A., Omar M., (2011) "Early Diagnosis of Diabetic Neuropathyin AlmadinahAlmunawwarah" Journal of Taibah University Medical Sciences 2011; 6(2): 121-134

[11] Oguejiofor O., Onwukwe C., Ezeude C., Okonkwo E., Nwalozie J., Odenigbo C., Oguejiofor C., (2019)."Peripheral neuropathy and its clinical correlates in Type 2 diabetic subjects without neuropathic symptoms in Nnewi, South-Eastern Nigeria". J Diabetol; 10:21-4.

[12] Reshma R, Anjana S Chandran (2020). "Literature Survey on Different Techniques Used for Predicting Diabetes Mellitus". IJCRT2004125 International Journal of Creative Research Thoughts (IJCRT) www.ijcrt.org

[13] Shahabeddin A., Sharareh R., Mehdi E., Hajar H., Ali G., (2019). "Artificial Intelligence Applications in Type 2 Diabetes Mellitus Care: Focus on Machine Learning Methods". Healthc Inform Res. 2019 October; 25(4):248-261. https://doi.org/10.4258/hir.2019.25.4.248 pISSN 2093-3681 • eISSN 2093-369X

[14] Thibault V, Belanger M, LeBlanc E, Babin L, Halpine S, Greene B, (2016)." Factors that Could Explain the Increasing Prevalence of Type 2 Diabetes among Adults in a Canadian Province: a Critical Review and Analysis". Diabetology & Metabolic Syndrome. 2016;8:71

[15] Wayne B. (2010) "Continuous glucose monitoring: Real time algorithms for calibration, filtering and alarms"; Journal of diabetes science and technology; vol 4; Issue 2; pp 404-418

[16] Zhang W., Zhong J., Yang S., Gao Z., Hu J., Chen Y., & Yi Z. (2019). Automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowledge-Based Systems,* 175, 12-25.