CLASSIFICATION OF DEMENTIA DISEASES USING DEEP LEARNING TECHNIQUES

[1]David O. Oyewola, [2]Bernard Alechenu, [3]Kuluwa A. Al-Mustapha, [4]Oluwatoyosi V. Oyewande

[1,2]Department of Mathematics & Computer Science, Federal University Kashere, Gombe, Nigeria
[3]Department of Mathematics, Baze University, Abuja, Nigeria
[4]Department of Computer Science, Faculty of Sciences, National Open University of Nigeria

*Corresponding author's email: davidakaprof01@yahoo.com

ABSTRACT

Dementia is the most frequent degenerative sickness in adults where early diagnosis can forestall or prolong progression. In this study, we used a deep learning techniques for classification of dementia. Data were collected from OASIS database of all the patients receiving dementia screening. The data included the patient's sex, age, education, social economic status, Mini-Mental State Examination, Clinical Dementia Rating, Atlas Scaling Factor, Estimated Total Intracranial Volume and Normalized Whole Brain Volume. The performance of every algorithm is juxtaposed with Generalized Regression Neural Network (GRNN), Radial Basis Neural Network (RBNN), Multilayer Perceptron Neural Network (MPNN) and Long Short Term Memory (LSTM) using Sensitivity, Specificity, Detection Rate. The results show that with 100% efficiency, GRNN, RBNN and LSTM tend to be the best in the classification of dementia. The use of deep learning such as LSTM for early diagnosis of dementia can help improve the process of dementia diagnosis.

Keyword: Generalized Regression, Long Short Term Memory, Radial Basis, Dementia, Perceptron.

INTRODUCTION

Approximately 4.5 million Americans have Alzheimer's ailment (AD), the eighth leading reason of loss of life in 2001. About 12.8% of those over the age of sixty-five and possibly as many as thirty-five to forty percent of these over the age of eighty are affected. There is no statistics of the real causes of AD and there is no diagnosed remedy. AD is a revolutionary neuro degenerative ailment resulting in a gradual decline in someone's reminiscence and capacity to learn, think, make judgments, interact, and carry out everyday activities. By 2050, with a growing population of an aging populace combined with a rise in lifespan, an estimated 11.3 to 16 million Americans will be afflicted by Alzheimer's, whose period can be among and greater than two decades. Former President Reagan's diagnosis and his death brought greater interest to the disease and its results on patients and their families. The brief discusses what is understood about the ailment, its human and financial outcomes, promising research areas, as well as policy issues and considerations. Dementia is a psychiatric condition that consists of a loss of intellectual capability (Gilman, 2010). According to (Igor, 2014), which highlights the scientific and neurobiological dimensions of AD and mild cognitive impairment (MCI) ought to be familiar with scientific college students. The paper also outlines developments in the use of biomarkers for AD diagnosis and discusses ongoing efforts to broaden new treatment plans. A paper by Aram et al, 2017 introduced a two-layer approach for diagnosis of dementia and application of machine learning techniques for early diagnosis of dementia. The information accrued have been from patients who receive treatment at Dementia Center in the Republic of Korea who provided dementia screening from 2008 to 2013. The authors utilized Precision, Recall and F-measure in juxtaposed other algorithms with Naive Bayes, Bayes Network, Logistic Regression, Random Forest, Support Vector Machine (SVM) and Multilayer Perceptron (MLP). Comparing dementia, mild cognitive impairment, normal and F-measurement values, the MLP achieved lowest, F-measurement values with 0.97, while the MLP was the highest in normal F-measurement values, while in MCI and dementia the SVM appears to be the highest with 0.739. The result suggests that the proposed dementia diagnostic model lessen the time and financial hardship and can offer assistance to improve the dementia diagnostic process. According to (Chen and Herskovits, 2010) in their evaluation of different statistical and machine learning approaches Naive Bayes and Support Vector Machine performed when compared with other algorithms. Machine learning and neural network approaches have been utilized in a research by Joshi et al. According to (Joshi et. al, 2009) they improve the accuracy of dementia screening devices, MMSE and Questionnaire Functional Activities. Their results showed that by integrating neural network with machine learning, the accuracy can be improved. In a paper by (Williams and Weakley, 2013) they apply Naive Bayes, Decision Tree, Neural Network, and SVM in dementia screening. Their findings show that Naive Bayes performs better than SVM. According to (Cho and Chen, 2012) suggested a double-layer hierarchical structure for early dementia diagnosis. This is a model that predicts early dementia diagnosis utilizing a top-layer Bayesian network following diagnostic prediction with FCM and PNN algorithm in the base layer while carrying out a psychological test which incorporate MMSE and CERAD. In this model, the accuracy of FCM and PNN was 74% and 69% respectively, however when contrasted with, MCI, and dementia, MCI and dementia were no longer well described. According to (Shanklea et. al, 1998) utilized machine learning techniques and electronic medical records to perform CDR prediction. The accuracy was the highest for

Naive Bayes, despite the fact it was lower than Bayesian for other algorithms, however it was about 70% accurate.

METHODOLOGY
In this study, the OASIS Database is a quantitative neuroimaging, clinical, cognitive and biomarker dataset for normal aging and Alzheimer's disease with 149 cases, 2 groups (dementia and non-dementia) and 9 attributes such as: age, sex, education, SES, MMSE, CDR, ASF, eTIV and nWBV. Our approach involves the use of deep learning techniques such as: Generalized Neural Regression Network (GRNN), Radial Neural Base Network (RBNN), Multilayer Perceptron Neural Network (MPNN) and Long Short Term Memory (LSTM)

Table 1: Longitudinal MRI Attributes

Terms	Meaning
Age	Age at time of image acquisition
Sex	Male or Female
Education	Years of Education
SES	Socio Economic Status
MMSE	Mini-Mental State Examination Score
CDR	Clinical Dementia Rating
ASF	Atlas Scaling Factor
eTIV	Estimated Total Intracranial Volume (cm^3)
nWBV	Normalized Whole Brain Volume

**Generalized Regression Neural Network (GRNN)**
A GRNN is a variant of the neural radial basis networks based on the Networks of kernel regression (Celikoglu et. al, 2007; Cigizoglu, 2005). As back propagation networks, a GRNN does not need an iterative training process. A generalized neural regression network (GRNN) is utilized for function approximation, as indicated by Dvořáková and Vochozka (2015). It has a radial layer of establishment and a particular longitudinal layer. All GRNNs have four layers (layer of input, hidden layer, layer of summation, and layer of output). The most disadvantages of GRNN models is their size and longer computation time during scoring compared to multilayer perceptron networks as there is one neuron per training row (Vochozka and Sheng, 2016) as shown in Figure 1.

The basic equation for GRNN is:

$$Y = \frac{\sum_{i=1}^{n} y_i e^{-\gamma}}{\sum_{i=1}^{n} e^{-\gamma}} \tag{1}$$

$$\gamma = \sum_{k=1}^{m} \left(\frac{x_i - x_{ik}}{\sigma}\right)^2 \tag{2}$$

Where y_i is the weight connection between the ith neuron, *n* is the number of the input training, *m* is the number of elements of an input vector, x_k and x_{ik} are the *jth* element of *x* and x_i, respectively, σ is the spread parameter, whose optimal value is experimentally determined.
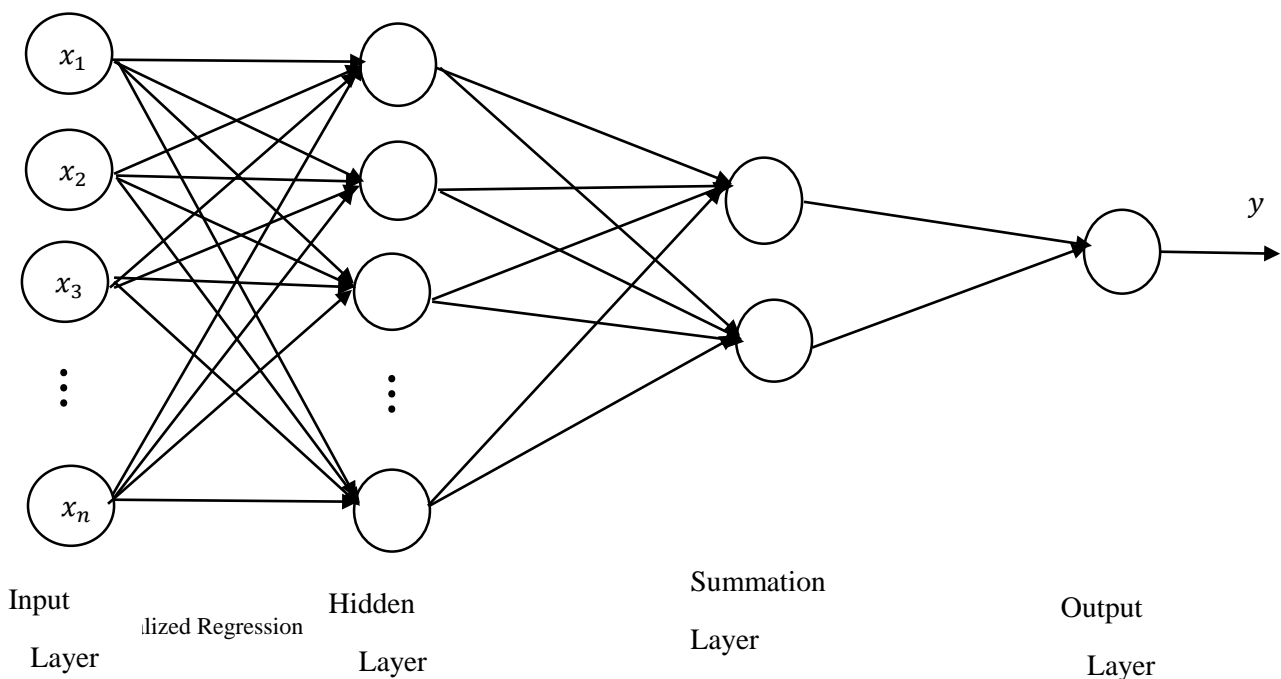
Radial Basis Neural Network (RBNN)

Radial basis functions are integrated into a neural network with two layers of feed-forward. A lot of inputs and a lot of outputs characterize such a network. Hidden layers is situated in between the inputs and outputs layers. A radial basis function is implemented on every one of the layers. The structure of a standard RBNN consists of three layers, to be specific the input, hidden and the output layers. There are few neurons in every one of these three layers. The input variables are equivalent to the input layer which consists of number of neurons and also input variables is equivalent to input layer neurons. The output layer neurons perform the outcome of the problem. Different outputs from an RBNN model can be executed. In any case, in this research a single neuron is considered for the output layer. Unlike the input and output layers, it was not easy to determine the quantity of hidden layer neurons and the proper measure of neurons for each problem is different. The RBNN are added to the input variables after the hidden layer moved to the neurons of the input layer. Hidden layer could be known as the RBNN's computational center gathering the input neurons utilizing the weighted summation and moving them through radial-basis functions to the nonlinear future (Chen et al. 2014).

For RBNN networks, the Gaussian activation function is given by:

$$\varphi_j(x) = e^{[-(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)]} \tag{3}$$

Where x is the input variable, μ_j is the mean and $\sum_j$ the covariance matrix of the jth Gaussian function.

The output layer implements a weighted sum of hidden-unit outputs:

$$\alpha_k(x) = \sum_{j=1}^{L} \tau_{jk}\varphi_j(x) \tag{4}$$

Where τ_{jk} are the output weights, each corresponding to the connection between a hidden unit and an output unit and L is the number of hidden units (Figure 2).

In classification applications, a sigmoidal function restricts the performance of the radial basis function to the interval (0,1):

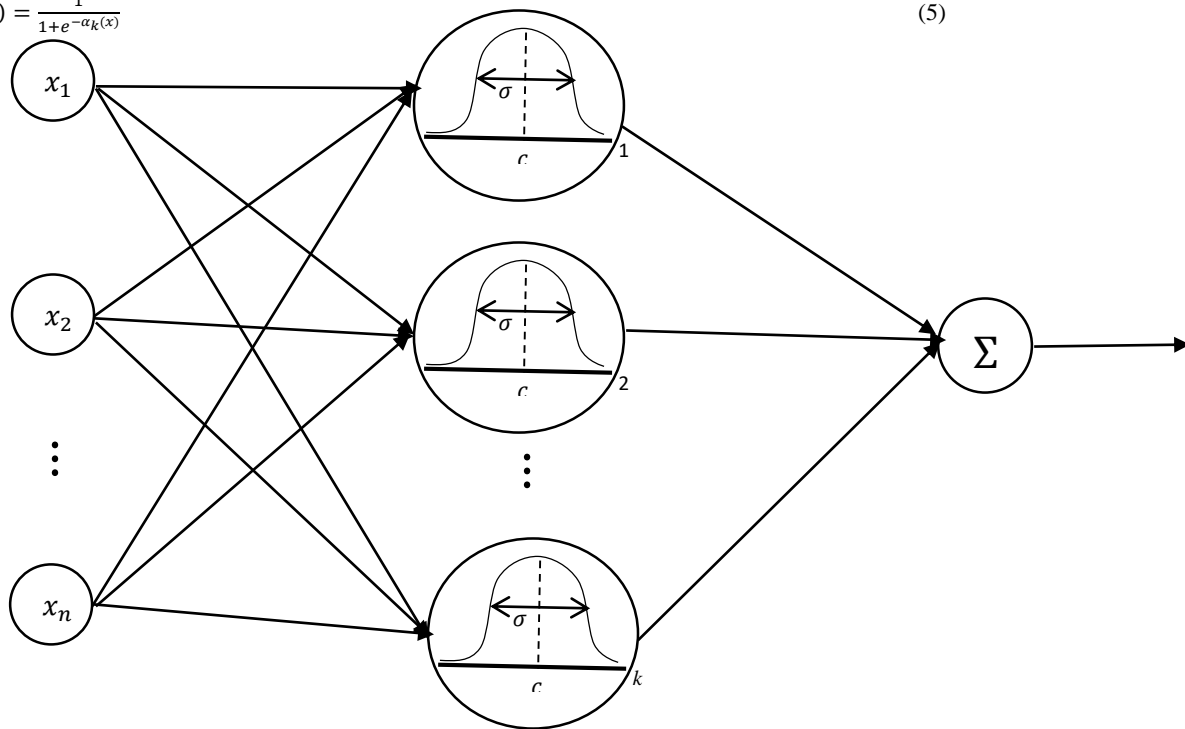$$Y(x) = \frac{1}{1+e^{-\alpha_k(x)}} \tag{5}$$

Figure 2: Radial Basis Neural Network

Multilayer Perceptron Neural Network (MPNN)

Figure 5 display the structure of an MLPNN which consists of three layers such as input layer with numerous neurons, an output layer where perceptron response is accessible with numerous neurons and hidden layer or intermediate layer. Although an MLPNN comprising of just input and output layers gives fulfilment for linear problems, it requires additional intermediate layers to estimate nonlinear problems. For instance, with just single hidden layer, all problems that can be solved by a perceptron, however utilizing (at least two) hidden layers is sometimes more effective (Dvorakova, 2015 and Vochozka, 2016).

The performance of the MLP NN is determined using:

$$n_1 = W_1 \times P + b_1 \tag{6}$$

where W is the matrix of the connection weight from the input nodes to the neurons of hidden layer, b_1 is the matrix of the neuron's bias, and P is the matrix of the input node. The output of each hidden layer neuron is calculated using a sigmoid function as shown in Eq. (7):

$$\psi = \frac{1}{1+e^{-n_1}} \tag{7}$$

Final outputs after calculating the hidden nodes output can be defined as:

$$n_2 = W_2 \times P + b_2 \tag{8}$$

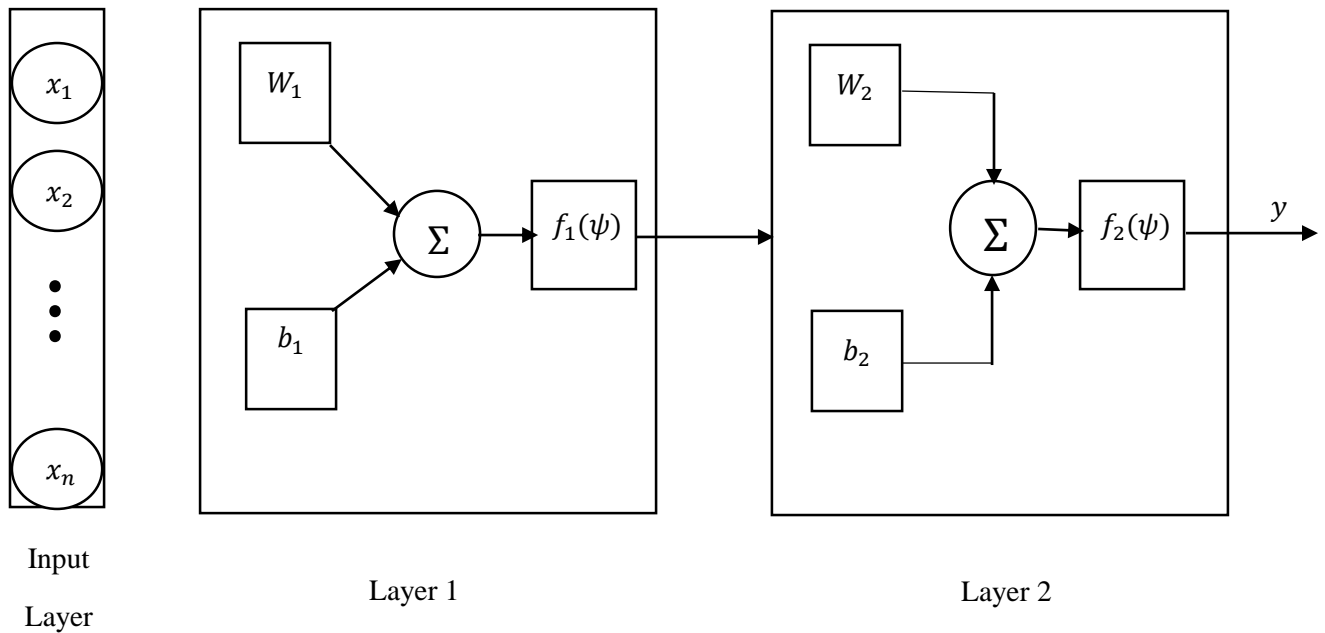$$Y = \frac{1}{1+e^{-n_2}} \tag{9}$$

Figure 3: Multi-Layer Perceptron Neural Network

Long Short Term Memory (LSTM)
LSTM is a variant of RNN developed by Hochreiter and Schmidhuber in 1997. The hidden layers are supplanted with LSTM cells on account of LSTM architecture. The cells comprise of various gates that can control the flow inputs. Every LSTM cell comprises of gate input, cell state, gate forgetting, and gate output. It likewise comprises of sigmoid layer, *tanh* layer and point wise multiplication process. The different gates and their functions are as follows:

- Input gate: Input gate comprises of the input.
- Cell State: Goes through the whole network and can utilize gates to include or remove data with the aid of gates.
- Forget gate layer: Chooses to permit a small amount of the data.
- Output gate: It comprises of LSTM output.
- Sigmoid layer produces zero to one, depicting the amount every component ought to be allowed to pass.
- Tanh layer produces a new vector, to add to the state.

The condition of the cell is changed based on the gates outputs.
We can describe it mathematically using the following equations:

$$f_t = \rho\big(W_f.[h_{t-1}, x_t] + b_f\big) \tag{10}$$

$$i_t = \rho(W_i.[h_{t-1}, x_t] + b_i) \tag{11}$$

$$c_t = tanh(W_i.[h_{t-1}, x_t] + b_c) \tag{12}$$

$$o_t = \rho(W_o.[h_{t-1}, x_t] + b_o) \tag{13}$$

$$h_t = o_t \times \tanh(c_t) \tag{14}$$

Where f_t forget gate vector, i_t input gate vector, c_t cell statevector, o_t outputgate vector, x_t is the input vector, h_t is the output vector and W, b are the parameter matrix and vector
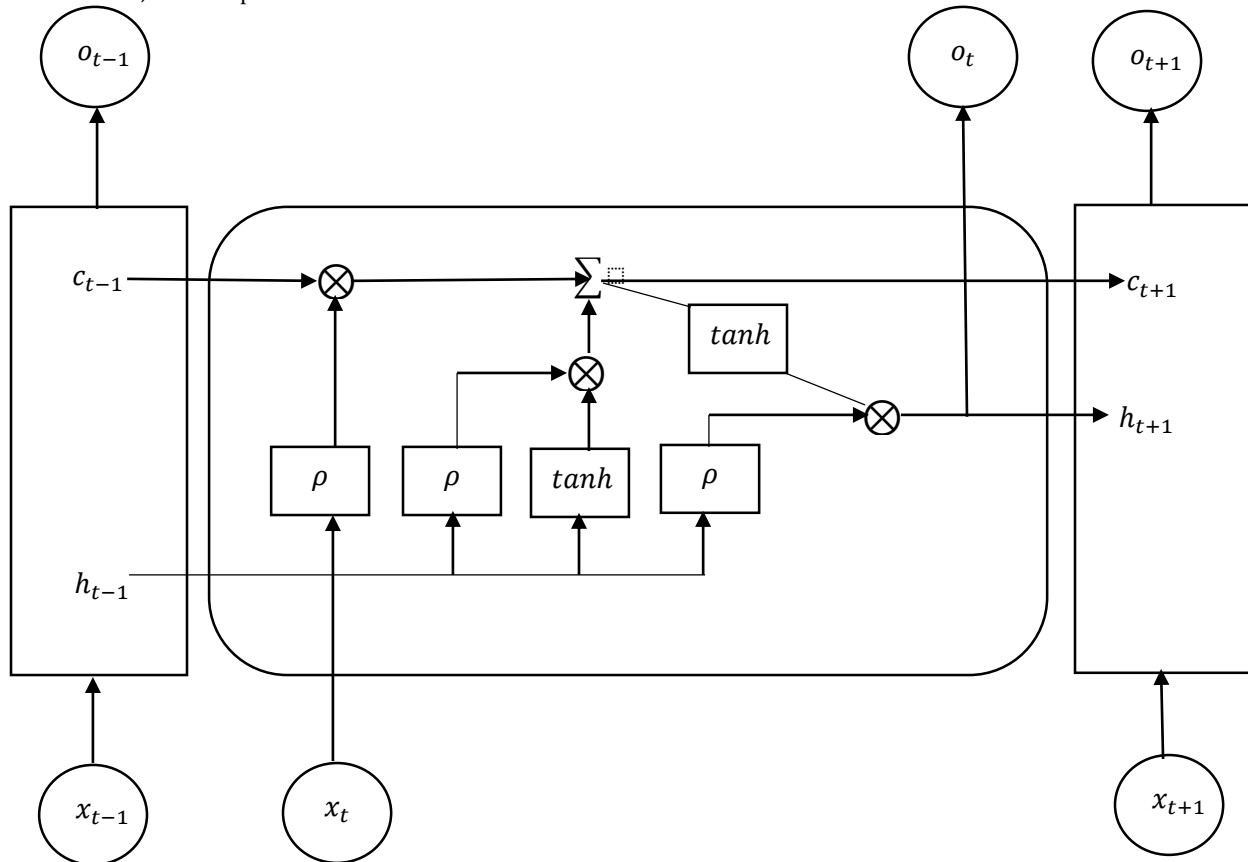

Figure 4: Long Short Term Memory

Performance Metrics

To determine the true value of a set data confusion matrix is utilized to evaluate a classification model's performance. The performance measures considered in this paper includes accuracy, sensitivity, specificity and detection rate determined from the confusion matrix.

Table 2: Confusion Matrix

Actual Class	Predicted Class	
	True Positive (TP)	False Positive (FP)
	False Negative (FN)	True Negative (TN)

True Positive(TP): This instance indicates dementia samples that were classified as dementia.
True Negative(TN): This instance indicates non-dementia samples that were classified as non-dementia.
False Positive(FP): This instance indicates dementia samples that were classified as non-dementia.
False Negative: It indicates non-dementia samples that were classified as dementia.
Performance metrics such as accuracy, sensitivity and specificity is the most widely used medicine and biology. The performance metrics are presented in Table 3.

Table 3: Performance Metrics formula

Performance Metrics	Formula
Accuracy	$\dfrac{TP + TN}{TP + TN + FP + FN}$
Sensitivity	$\dfrac{TP}{TP + FN}$

Specificity	$\dfrac{TN}{TN + FP}$
Detection Rate	$\dfrac{TP}{TP + TN + FP + FN}$

RESULT AND DISCUSSION

This section analyzes the results of the predictive dementia and non-dementia disorder using generalized regression, radial basis, multilayer perceptron, and long-term memory. Dementia-related and non-dementia-related data was obtained from 373 patients, while 149 patients have been identified. To compare visually the two groups such as dementia and non-dementia patients utilized in this research. Table 4 presents the cross tabulation of dementia and non-dementia condition by respondents Age, Sex, Education, SES, MMSE. The first column includes the variables: age, sex, gender, SES and MMSE as well as their categories, column two contain the percentage of dementia and the last column captured the non-dementia and its percentages. Table 4 results show that there is no change in the state of dementia when patients age while the level of non-dementia is correspondingly increased. Both male and female record higher dementia status rates. From Table 4, age group strength leans strongly towards older adults with a range of 54-56%. Most respondents are in the condition of non-dementia. The result shows that 60% of males have dementia, while 63% are female. Figure 1 is the missing map that shows no missing data.

Table 4: Cross Tabulation of Dementia and Non-Dementia Condition

| Variables		Dementia (%)	Non-Dementia (%)
Age	Adult(60-80)	114(45.60)	136(54.40)
	Older Adult (81-100)	53(43.09)	70(56.91)
Sex	Male	96(60)	64(40)
	Female	71(62.83)	42(37.17)
Education	Few Educ.(5-15)	107(52.71)	96(47.29)
	Long Educ. (16-25)	60(35.29)	110(64.71)
SES	Upper= 1	38(43.18)	50(56.82)
	Upper-Middle= 2	28(27.18)	75(72.82)
	Middle= 3	37(45.12)	45(54.88)
	Working= 4	40(54.05)	34(45.95)
	Lower =5	5(71.43)	2(28.57)
MMSE	Worst= 0	89(97.8)	2(2.20)
	Best ≥ 27	78(27.66)	204(72.34)
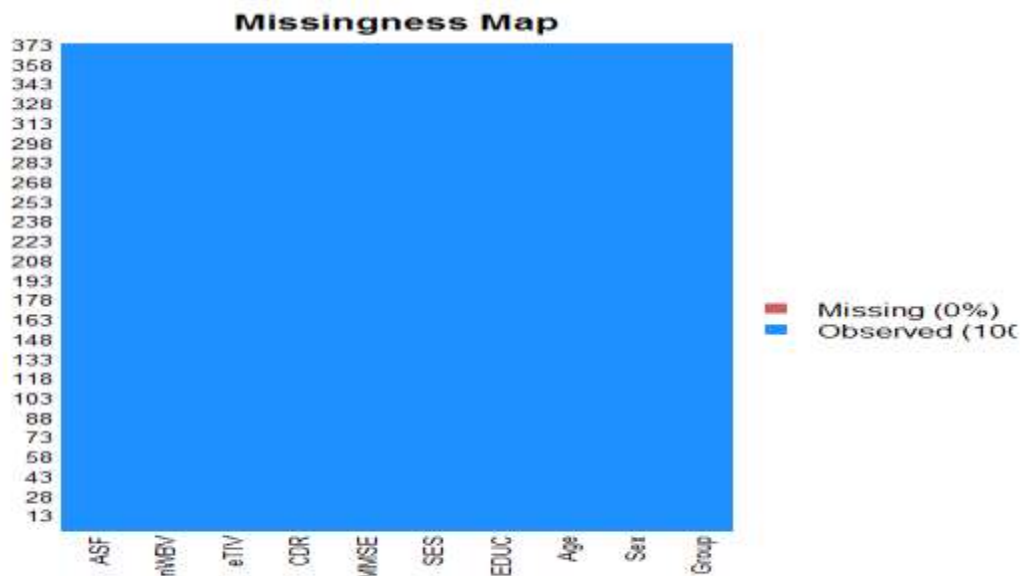
Figure 5: Missing Map

The early diagnosis of dementia, as explained above, is based on the results of deep learning techniques such as long-term memory (LSTM), radial neural basis (RBNN), multilayer perceptron neural network (MPNN) and generalized neural regression (GRNN). We juxtapose dementia and non-dementia diagnosis to obtain the more accurate algorithms in diagnosis the diseases. The correctly classified dementia diagnostic data was observed and its accuracy is measured as shown in Table 5 for the four deep learning. After the four deep learning classification model has been completed. Use 149 OASIS dataset clinical cases. Table 6 reveals that LSTM, RBNN and GRNN perform better at (100%) than MPNN (58.4%) in terms of accuracy than deep learning. The findings show that it is of great advantage to use deep learning methods to diagnose dementia. Nevertheless, motivated by the perspective utilized for early diagnosis in dementia treatment centers, our model can not only diagnose dementia with data from specific patient assessments, but it also aims to achieve greater precision in early dementia diagnosis. To organize and visualize the output obtained from classifier Receiver Operating Characteristics (ROC) is very useful as shown in Figure 5.ROC graphs has two axes the vertical axis represents the Y axis while the horizontal axis is called the X axis plotting true positive rate, false positive rate respectively. Figure 2 shows a ROC graph with four deep learning classifiers that were used on the same graph in this paper. The best curve was obtained from GRNN, RBNN and LSTM from Figure 6 above as it is closer to True positive rate to the northwest.

Table 5: Classification of Deep Learning

DL	Diagnosis	Dementia	Non-Dementia
LSTM	Dementia	70	0
	Non-Dementia	0	79
RBNN	Dementia	70	0
	Non-Dementia	0	79
MPNN	Dementia	62	8
	Non-Dementia	54	25
GRNN	Dementia	70	0
	Non-Dementia	0	79

Table 6: Performance Metrics

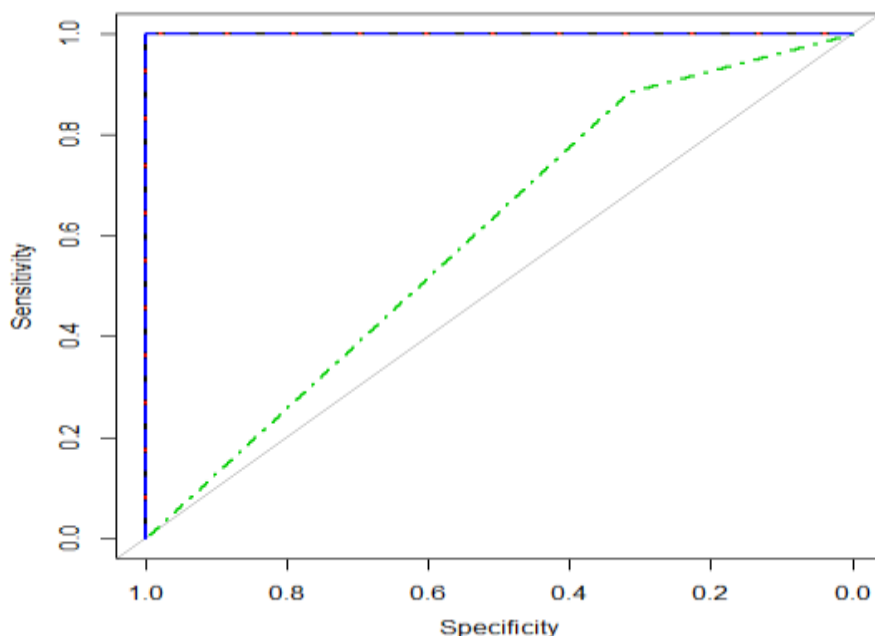| DL	Accuracy	Sensitivity	Specificity	Detection Rate
LSTM	1	1	1	0.4698
RBNN	1	1	1	0.4698
MPNN	0.5839	0.5345	0.7576	0.4161
GRNN	1	1	1	0.4698

Figure 6: ROC Curve

CONCLUSIONS

The main purpose of this study is to explore the capability and applicability of machine and deep learning methods such as GRNN, RBNN, MPNN and LSTM for classification of dementia disease. We used a machine and deep learning model in this study for early dementia diagnosis. OASIS data were utilized to minimize individual time and financial burden in screening with the aid of deep learning algorithms. We compared different machine and deep learning classification models using data on dementia diagnosis. The deep learning model demonstrate the effort of translating test results by creating a set of criteria for patient classification and therefore early-stage diagnosis of dementia in a simple, inexpensive and reliable manner that enhances current clinical practice. A deep learning model architecture is provided in this paper to distinguish dementia and non-dementia diseases. Confusion Matrix is used to test the model and class results (dementia and non-dementia). The experimental result shows that the accuracy of GRNN, LSTM, and RBNN is 100% higher than the norm of machine learning. For further analyses, other predictive models, such as Feed forward neural network, recurrent neural network, can be applied.

REFERENCES

Celikoglu HB, Cigizoglu HK. (2007). Public transportation trip flow modeling with generalized regression neural networks. Adv. Eng. Software 2007;38:71–9.

Cigizoglu HK, Alp M. (2005). Generalized regression neural network in modelling river sediment yield. Adv. Eng. Software 2005;37: 63–8.

Bors, A. G., Gabbouj, G., (1994) "Minimal topology for a radial basis function neural network for pattern classification," *Digital Signal Processing: a review journal*, vol. 4, no. 3, pp. 173-188.

Kim B, Lee DW, Parka KY, Choi SR, Choi S (2004). Prediction of plasma etching using a randomized generalized regression neural network. Vacuum 2004; 76:37–43.

Dvořáková, l. and M. Vochozka, (2015). Vykorystannjanejronnychmereždlja prognozuvannjarozvytkukompaniji. *Nacional'naakademijanaukUkrajinyKijiv*. (3), 5-12. ISSN 2409-8876.

Vochozka, M. and P. Sheng, (2016). The application of artificial neural networks on the prediction of the future financial development of transport companies. *Komunikácie*.**18**(2), 62-67.

Chen W, Fu ZJ, Chen CS (2014). Recent advances in radial basis function collocation methods. Springer, Berlin.

Gilman S (2010). Oxford American handbook of neurology. Oxford University Press: Oxford, UK; 2010.

Igor O. Korolev (2014). Alzheimer's Disease: A Clinical and Basic Science Review, Medical Student Research Journal, vol. 4 pp. 24-33.

Aram So, DanialHooshyar, Kun Woo Park and HeuiSeok Lim (2017), Early Diagnosis of Dementia from Clinical Data by Machine Learning Techniques, Appl. Sci. *2017*,1-17

Chen, R., Herskovits, E.H (2010). Machine-learning techniques for building a diagnostic model for very milddementia. *Neuroimage* 2010, *52*, 234–244.

Joshi, S., Shenoy, P.D., Venugopal, K.R., Patnaik L.M (2009), Evaluation of different stages of dementia employing neuropsychological and machine learning techniques. In Proceedings of the First International Conference on Advanced Computing, Chennai, India, 13–15 December 2009.

Williams, J.A, Weakley, A., Cook, D.J. and Schmitter-Edgecombe M. (2013), Machine learning techniques for diagnostic differentiation of mild cognitive impairment and dementia. In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), Bellevue, WA, USA, 14–18 July 2013.

Cho, P.-C., Chen, W.-H (2012). A double layer dementia diagnosis system using machine Learning techniques. In Proceedings of the 13th International Conference, (EANN 2012), London, UK, 20–23, September 2012.

Shanklea, W.R., Mania, S., Dick, M.B., Pazzani, M.J (1998). Simple models for estimating dementia severity using machine learning. *Stud. Health Technol. Inform.* 1998**,***52 Pt 1*, 472–476.

Shaikh Abdul Hannan, R. R. Manza, R. J. Ramteke, (2010). Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis, International Journal of Computer Applications, vol. 7 (13) 7-13.

David Oyewola, Danladi Hakimi, Kayode Adeboye, Musa Danjuma Shehu (2016). Using Five Machine Learning for Breast Cancer Biopsy Predictions Based on Mammographic Diagnosis, International Journal of Engineering Technologies-Ijet, Vol.2, No.4, 142-145.