Artificial Neural Networks: Its Techniques and Applications to Forecasting

Anish Gupta
Assistant Dean Academics,
Amity University,
Greater Noida, India
agupta4@gn.amity.edu

Ayodeji Olalekan Salau
Dept. of Electrical/ Electronics and Computer Engineering,
Afe Babalola University,
Ado-Ekiti, Nigeria
ayodejisalau98@gmail.com

Prateek Chaturvedi
Dept. of Mechanical Engineering,
Amity University,
India
pchaturvedi1@gn.amity.edu

Segun Ayokunle Akinola
Dept. of Electrical Electronics Engineering,
University of Johannesburg,
South Africa
akinolasegun35@yahoo.com

Nnamdi Ikechi Nwulu
Dept. of Electrical Electronics Engineering,
University of Johannesburg,
South Africa
nnwul@j.ac.za

***Abstract*—In the last few decades, Artificial Intelligence (AI) techniques have been used extensively in numerous applications. Amongst AIs numerous techniques, the Artificial Neural Networks (ANN) has dominated in its capability and applicability in different fields. Neural is a word derived from (animal) human nervous system which represents the nerve cells or neurons present in the brain. This paper however focuses on the review of ANN and its applications in forecasting. The basic concepts of ANN, its characteristics and applications were discussed. In addition, a comparison was made between ANN and other conventional methods. The conclusion of this review shows that ANN mostly performs better than other conventional techniques in solving real-life problems.**

Keywords—Artificial neural network, techniques, applications, data

I. Introduction

Over the years, Artificial Neural Networks (ANN) have been a robust and capable technique for solving real-life problems in many fields of study such as medicine, finance, industry and in various applications of engineering. ANN works like the human brain and has the capacity for learning the patterns and signals of input data before converting to a corresponding output. This technique has a robust capacity for forecasting, classification, and prediction of data [1].

ANN is a data-driven technique which belongs to the family of Artificial Intelligence (AI). It is a tool used for non-linear estimation, sorting of data, optimization, pattern detection, and for clustering. In addition, it is a very robust tool for modelling and simulation. Even when the underlying data relationship is unknown, ANN still learns and identifies correlated patterns among the inputted data set with the output targets during training. This is with the aim of getting an independence outcome of the inputted data set. It has the ability to process non-linear, complex and noisy datasets. It is also known for its fast processing, high accuracy and speed [1].

ANN has the capability of accurately predicting the outcome of a dataset, even when it has very little and otherwise incomplete knowledge of the problem. This is a major strength of the ANN technique [2]. A network is said to be neutral if it is inspired by the neuroscience theory which involves learning of patterns and other cognitive processes. It involves some statistical and mathematical processes such as clustering, statistical regression, non-parametric pattern classification and non-filter models which are neurobiological models.

Similar to the conventional statistical techniques, ANN has a very wide range of application such as in classification problems, speech recognition, identification of under-water sonar currents, forecasting, time series, prediction of secondary structure in globular proteins and many more [3]. Statistical problems can be solved through classical statistical techniques such as Bayes analysis, discriminate analysis, multiple regression, logistic regression and ARIMA time series [3].

II. Applications and Functions of ANN

A. Meaning of Neural Network

Neural is a word derived from (animal) human nervous system which represents the nerve cells or neurons present in the brain. A picture representation of a neuron is shown in Fig. 1.

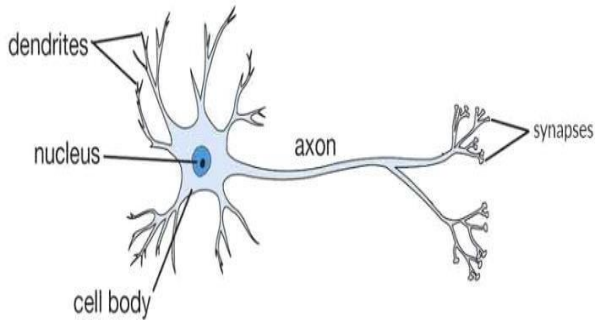

Fig. 1. Biological Neuron [4].

B. Neuron parts and their functions

There are four parts of the human brain nerve cells namely:

- Dendrites:- Signal is received from other neurons.
- Cell body:- It adds up all the incoming signals to

generate the input.

- Synapses:- It is the interconnection between one neuron and the other. The signal amount is dependent on the strength of synaptic weight of connections
- Axon:- At the point, the sum of neuron signals gets to a threshold value where the neuron fires with the signal traveling down the axon to other neurons. Connections which are inhibitory naturally decrease in strength or increase in strength.

Generally, Neural Networks (NN) is an interconnected network of billions of neurons with a lot of interconnections among them [1]. The differences between the computer and human brain is presented in Fig. 2 and an architecture of NN is shown in Fig. 3.

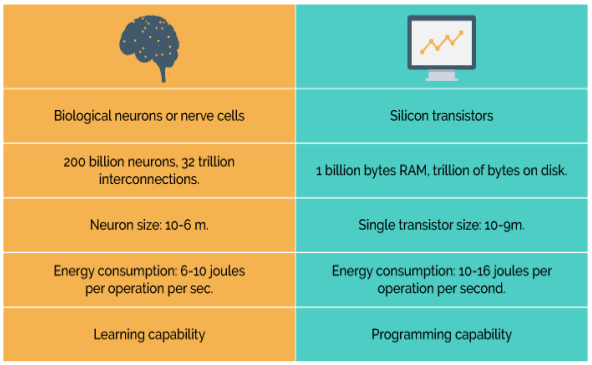

Fig. 2. Difference between the Computer and Human Brain [4].

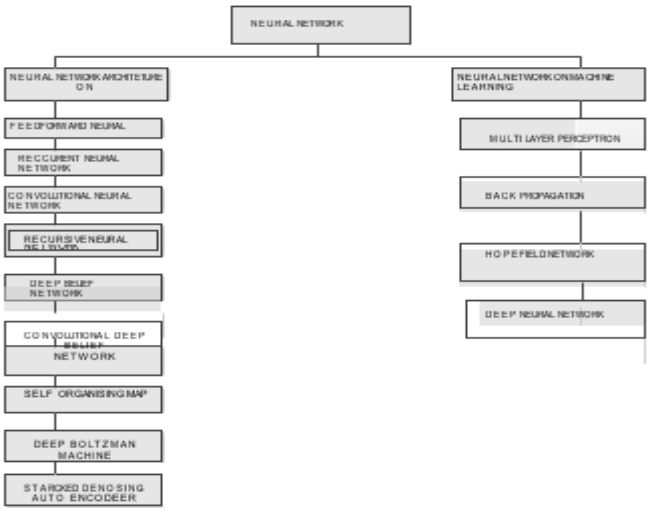

Fig. 3. Architecture of NN.

C. *Neural Network Characteristics*

- The performance of NN is very excellent in mapping out the input data with the pattern associated with it.
- In a very high speed, NN will process information in parallel in a distributing manner.
- NN is very good in tolerating a faulty system which will make it recall the pattern from the incomplete noisy or partial pattern.
- NN has good skills in pattern identification and learning from previous examples during training.

Although, ANN is derived from the concept of biological NN, they have a number of differences as presented in Table I.

TABLE I. DIFFERENCE BETWEEN ARTIFICIAL NEURAL NETWORK AND BIOLOGICAL NEURAL NETWORK

Characteristics	Artificial Neural Network	Biological Neural Network
Storage	Information stored recently added can delete the old data	A very high complexity and dense network with interconnected neurons which contain neuron from 1011 to 1015
Size and Complexity	It has less size and complexity and does not perform pattern recognition task	It has high complex and dense network with interconnected neurons that contain neurons in order of 1011 and 1015 of interconnection
Speed	Very fast in information processing using less than seconds	A bit slower in processing information with milliseconds
Control Mechanism	It has a control unit controlling all the activities	No main control unit but external to computer disk
Processing	Processing is serial	Processing is parallel
Fault tolerance	Corrupt information cannot be retrieved when the system has failed	The stored information is adaptable to new information added with the system still keeping the old information

III. TRAINING AND LEARNING METHODS

The learning methods adopted in an NN are basically classified into three. These are unsupervised, supervised, and reinforced learning. Each of them are discussed in (A), (B) and (C).

A. *Unsupervised learning*

The output, which is the target is not presented in this leaning technique. In an unsupervised learning, it is assumed that there is no teacher which will present the desired pattern, hence, the system has to learn by itself by adapting and discovering the features and structure in the input patterns.

B. *Supervised learning*

In supervised learning, the input is used to train the network which is associated with the output pattern and the target or desired pattern. The teacher is assumed to be present during the learning process. During a comparison between the network that is computed, the correct expected output determines the error which can be used in changing network parameters that will help to improve performance.

C. *Reinforced learning*

In reinforced learning, the teacher, though available, will not present the expected answer but will be indicated when it is computed and if incorrect or correct, while the information provided helps the network to learn in the process. The correct answer is rewarded while there is a penalty for a wrong answer. Among all the learning types, reinforced learning is the least effective.

IV. TYPE OF NEURAL NETWORKS

A. *Dynamic*

In dynamic NN, the output always depends on the current input (memory unit). An example of this is Recurrent Neural Network (RNN).

B. *Static*

In static NN, the output also depends on the current input (memory-less unit). An example is feedforward network.

C. *Multilayer*

This consist of multiple hidden layers within the network. An example is the multilayer perceptron.

D. *Feedforward*

This is where the graphs have no loops.

E. *Fixed*

The weight is fixed a priori and not changed in any way.

F. *Adaptive*

The weights are updated and changed during the process of training.

V. USES OF NEURAL NETWORK TECHNIQUES

A. *Classification*

The NN is trained to classify the given pattern or data set to pre-defined class.

B. *Predicition*

The NN is trained to produce output that is expected in the given input.

C. *Clustering*

The NN is used in the identification of some important features of the data when classifying them into different categories without the knowledge of the data.

D. *Association*

The NN is trained for remembering some patterns so that the noise pattern is presented to the network and the network will associate with the closest one in the memory.

There are numerous advantages and disadvantages of NN; some are presented in Table II.

TABLE II. THE ADVANTAGES AND DISADVANTAGES OF NEURAL NETWORK

Advantages of Neural Network	**Disadvantages of Neural Network**
Neural Network will perform a task of a linear program	The neural network needs training generally to operate
Neural Network will perform a task without any problem	Require high processing time for large neural networks
When element fail still it will continue without problem with parallel nature	The architecture is different from the microprocessor

VI. APPLICATION OF NN TO FORECASTING

There are different research studies that have applied NN to solve real world problems in the past. Paramount among such is a study carried out in [5]. This study, developed a Multi-Layer Perceptron (MLP) model to forecast temperature at a minimum and maximum ground level in Dum Dum station. This forecast is based on the average daily data of some parameters such as vapor pressure, rainfall, level of pressure and radiation between 1989 and 1995. The use of the NN technique was also demonstrated in [6], [7]. The technique was applied to forecast plant disease. A NN model was developed for predicting financial status of the economy using time series. A model was developed to predict milk production in [8] by adopting a feedforward NN. In another study, a comparison was made between the performance of MLP and the linear regression for epidemiological data regarding the quality of prediction with the robustness to the deviation from underlying assumptions in normal homoscedasticity with the independence of error. A comparison of the performance of epidemiological data using various methods was presented in [9]. Feedforward NN was used to compare ARIMA with NN as a duo, were used for price prediction in [10]. Authors in [11], adopted ANN in estimating daily grass reference crop evapotranspiration. Furthermore, in the study, a comparison between ANN and conventional techniques was made with regards to their performance.

A set of guidelines for selecting the best ANN model and the advantages over other conventional statistical methods was highlighted in [12]. In the medical field of research, ANN was adopted to predict the Human Immunodeficiency Virus (HIV) drug resistance mutation among some patients who were infected with HIV. The result of the study shows that the ANN model has the capacity to identify both the effectively treated patients and those with failing therapies [13]. Furthermore, [14] conducted a study by adopting three computational techniques in predicting virological response which combines HIV therapy and treatment. A comparison of the performance accuracy was done between Radom Forest (RF) and Support Vector Machine (SVM). The result showed that both RF and SVM can produce good and accurate results. In a study conducted by [15], the ANN technique was adopted to predict HIV protease cleavage site protein. It was reported that the ANN technique produced a 92.06% predictability and this was helpful in finding the HIV protease inhibitor. In the study, a comparison of five conventional methods with ANN was made using sixteen input variables. The result of the performance shows that ANN performed better than other techniques and hence, it was recommended as an alternative to established traditional methods. Authors in [25], have used ANN and IoT for flood forecasting. Furthermore authors in [26] have made a review of IoT and its use in engineering applications. Among the recent trends in research is the use of ANN and machine learning techniques in computer vision applications, especially for pattern recognition and license plate detection [27].

VII. OVERVIEW OF NN TO DATA INTENSIVE APPLICATION

Table III presents some NN applications to various forms of data.

TABLE III. NEURAL NETWORK TO DATA APPLICATIONS

Architecture	Application	Activation function
Deep neural network	Searching intelligent	Logistic function
Modular neural network	Target recognition	Tan-Sigmoid
Back propagation algorithm	Marketing	Logistic function
Multi-later perceptron	Medical	Tan sigmoid
Classification supervised	Management portfolio	Tan sigmoid

algorithm		
Radial basic function	Modelling and control	Radial basis
Convolutional neural network (CNN)	Voice and image recognition	Logistic function
Support vector machine	Rating credit	Logistic function
Back propagation	Forecasting	Logistic function

VIII. Designing A Neural Network Model

A. Selection of input variables

Selection of the input variables is one of the key steps in NN modelling. It has a great influence on the performance of the model. However, there are two basic factors that must be considered which are independence and significance of the input variables. [16] presented some strategies or methods for evaluating the similarities between the input and output parameters.

The model-free method simply involves analytical methods based on the inter-relationship and collective agreement and the ad-hoc manner in accordance to data visibility or knowledge section [17]. The stepwise method is the pruning, constructive sensitive analysis ad-hoc which is trial by error [18]. The optimization method is grouped into two parts namely, filtering and dimensionality reduction. The dimensionality reduction is simply accomplished with spinning the input or another way to cluster the data while the filtering technique is partially located by the agreed information or rather partial correlation.

B. Division of data

In the modelling process of NN, the datasets are divided into three subsets namely: the training, the testing, and the validation datasets. Estimating the unknown connection weight is achieved by the training dataset in order to get the attained optimal structure and avoid possible over fitting. The validation set will simply be used to evaluate the general performance of the trained model.

C. Selection of the model type

The selection of model type is another key factor to getting the optimal model performance of an NN architecture. NN model architecture includes Feed-forward, Recurrent, Hybrid neural network, radial basic function, multi-layer perceptron, probabilistic neural network (PNN), associative memory network, generalized regression neural network, and reformulated neural network [19]. Information goes from input to the output layer in one direction in the feed forward neural network, and multilayer perceptron is the commonly used structure for feedforward NN. In the RNN, the input data is passed to the output and it reverses back to the input through the feedback loop layers. The hybrid methods such as the ANFIS, RNN, Genetic Algorithm Aiding Artificial Neural Network (GANN) are used for optimal data prediction. Meanwhile, the RNN input passes the information to the output with a reverse back to the input through the feedback loop layer [20].

D. Model structure selection

In modelling NN, the number of nodes and layers are used to develop the model structure. This optimal structure reduces error predicted or gives an optimized forecast of the NN. Therefore, the NN structure that has the fewer number of hidden neurons are selected based on trial and error. No specific rule exists about the specific neurons that will be put in a hidden layer to get the performance. Although, there are some rules of the thumb to calculate the hidden neurons [21].

A lot of techniques can be used to determine the neural structure. Classified methods can get optimal ANN models though the regularization approach which can also be used in finding the optimal model of ANN. It is important to know that no method is required in the determination of the structure of the hybrid NN model because their structures are fixed [22].

E. Data normalization

Normalization of data is a process of restructuring a relational database with the aim of avoiding redundancy of the data and it is very important because of the variable units which are different in processing and training with the neural model which is efficient when using the training function in NN [23].

The main challenge of optimal prediction is the training of the calibration i.e. the model finding the optimal sets in variables that indicate the input and output data links which use the optimization learning algorithm. This method is normally used to determine the techniques, optimal set of variables reducing the error among the actual with predicted data. The deterministic techniques are grouped into three sets which are local, global and other techniques. The training methods which include: gradient descent, Levenberg Marquardt (LM), Jacobian matrix and Quasi-Newton back propagation techniques are used in getting standard and perfect results. Meanwhile, stochastic methods are also used for calibrating the model, though the Bayesian methods are generally used [22].

F. Training of data

The data will be grouped into two methods which are the supervised and unsupervised methods the randomize data splitting to physical base domain knowledge, the organizing map itself, then the ad-hoc methods are categorized as unsupervised data. These are generally used as random division data though the try and error method is the genetic algorithm with based optimization technique which is used in the supervised data division [20].

G. Evaluating ANN network models

After training the model, there is a need to evaluate the model performance by the use of statistical techniques. Popular among such is the statistical methods which measure the correlation coefficient (R), the coefficient of determination (R2), Mean Square Error (MSE), Mean Absolute Error (MAE), Percentage Error and Root Mean Square error (RMSE). These methods are used to measure the amount of error between the measured and predicted models.

Models are efficiently validated by the three aspects like replicative validity, structure validity, and predictive validity. Replicate valid is used only if the data matches with the original of the real system using the previous step of ANN structure modelling. The replicative validity is performed with statistical techniques, correlation analysis, a nova

analysis, goodness, and regression analysis by testing the model. The model represents the correlation of input with the output data. Meanwhile generalization is the capacity of the model to be examined by predictive validation to test above a ranging data use in calibrating. ANN models work by predicting the quantitative error measure. Techniques such as SSE, RE, MAE can be used for visual inspection to check ANN model performance. Structural validity test is used in analyzing the uncertainty of ANN output and the structural validity of the ANN examined when using different methods in connection with the sensitivity analysis, the degree off generalization, in comparing modelling results and the previous information modelling systems [22], [24].

IX. Conclusion

In the world of computing, ANN has performed excellently and has greatly been able to learn from example. This singular feature has made ANN a robust tool for solving real-life problems in many fields of study. A number of advantages and disadvantages of ANN have been highlighted in this paper. In some cases, ANN will not serve as a replacement to conventional methods, however, with the growing lists of applications, ANN has proved to be a good alternative to most conventional techniques in the world of computing. Also, it has been proved that ANN performs better when combined with conventional techniques for forecasting operations, pattern recognition and for solving general engineering problems.

References

[1] M. H. Hassoun, *Fundamentals of artificial neural networks*: MIT press, 1995.

[2] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the national academy of sciences*, vol. 79, pp. 2554-2558, 1982.

[3] B. Cheng and D. M. Titterington, "Neural networks: A review from a statistical perspective," *Statistical science*, pp. 2-30, 1994.

[4] N. S. Gill, "Artificial neural network application," *xenonstack*, 2017.

[5] S. Pal, J. Das, P. Sengupta and S. K. Banerjee, "Short-term prediction of atmospheric temperature using neural networks," Mausam, vol. 53, no. 4, pp. 471-480, 2002.

[6] E. De Wolf and L. France, "Neural network classification of tanspot and Stagonospora blotch infection periods in a wheat field environment," *Phytopathology*," vol. 90, pp. 108-113, 2000.

[7] I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, pp. 215-236, 1996.

[8] L. Sanzogni and D. Kerr, "Milk production estimates using feed forward artificial neural networks," *Computers and Electronics in Agriculture*, vol. 32, pp. 21-30, 2001.

[9] J. Gaudart, L. Huiart and B. Giusiano, "Comparison of the performance of multi- layer perceptron and linear regression for epidemiological data," *Computational Statistics and Data Analysis*, vol. 44, pp. 547-570, 2004.

[10] N. Kohzadi, M. S. Boyd, B. Kermanshahi and I. Kaastra, "A comparison of artificial neural network and time series models for forecasting commodity prices," *Neurocomputing*, vol. 10, pp. 169-181, 1996.

[11] M. Kumar, N. S. Raghuwanshi, R. Singh, W. W. Wallender, W. O. Pruitt, "Estimating evapotranspiration using the artificial neural network," *Journal of Irrigation and Drainage Engineering*, vol. 128, pp. 224-233, 2002.

[12] G. Zhang, B. E. Patuwo and Y. M. Hu, "Forecasting with artificial neural networks: The state of the art," *International Journal of forecasting*, vol. 14, pp. 35-62, 1998.

[13] A. D. Revell, D. Wang, M A. Boyd, S. Emery, A. L. Pozniak, F. De Wolf , "The development of an expert system to predict virological response to HIV therapy as part of an online treatment support tool," *Aids*, vol. 25, pp. 1855-1863, 2011.

[14] D. Moore, P. Aveyard, M. Connock, D. Wang, A. Fry-Smith and P. Barton, "Effectiveness and safety of nicotine replacement therapy assisted reduction to stop smoking: systematic review and meta-analysis," *British Medical Journal*, vol. 338, pp. 867-871, 2009.

[15] M. Toyota, "CpG island methylator phenotype in colorectal cancer," *Proceedings of the National Academy of Sciences*, vol. 96, pp. 8681-8686, 1999.

[16] G. J. Bowden, G. C. Dandy, and H. R. Maier, "Input determination for neural network models in water resources applications. Part 1—background and methodology," *Journal of Hydrology*, vol. 301, pp. 75-92, 2005.

[17] A. Sharma, "Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1—A strategy for system predictor identification," *Journal of Hydrology*, vol. 239, pp. 232-239, 2000.

[18] *M. J. Diamantopoulou*, V. Z. Antonopoulos and D. M. Papamichail, "Cascade correlation artificial neural networks for estimating missing monthly values of water quality parameters in rivers," *Water resources management*, vol. 21, pp. 649-662, 2007.

[19] A. Castelletti, S. Galelli, M. Restelli and R. Soncini-Sessa, "Data-driven dynamic emulation modelling for the optimal management of environmental systems," *Environmental modelling and software*, vol. 34, pp. 30-43, 2012.

[20] H. R. Maier, G. C. Dandy, A. Jain and K. P. Sudheer, "Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions," *Environmental modelling and software*, vol. 25, no. 8, pp. 891-909, 2010.

[21] S. M. Zahraee, M. Khalaji Assadi, R. Saidur, "Application of artificial intelligence methods for hybrid energy system optimization," *Renewable and sustainable energy reviews*, vol. 66, pp. 617-630, 2016.

[22] W. Wu, H. R. Maier and G. C. Dandy, "Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling," *Environmental Modelling and Software*, vol. 54, pp. 108-127, 2014.

[23] V. Sivamadhavi and R. S. Selvaraj, "Prediction of monthly means solar radiation using the artificial neural network," *J.Earth Syst Sci*, pp. 1501-1510, 2012.

[24] A. Jain and S. Kumar, "Dissection of trained neural network hydro logic models for knowledge extraction," *Water resources research*, vol. 45, no. 7, 2009. DOI: 10.1029/2008WR007194

[25] P. Mitras, R. Ray, R. Chatterjee and R. Basu, "Flood forecasting using Internet of things and artificial neural networks," *IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference*, Vancouver, BC, pp. 1-5, 2016. DOI: 10.1109/IEMCON.2016.7746363

[26] A. Kumar, A. O. Salau, S. Gupta, K. Paliwal, Recent trends in IoT and its requisition with IoT built engineering: A review. In: Rawat, B., Trivedi, A., Manhas, S., Karwal, V. (eds) Advances in Signal Processing and Communication. Lecture Notes in Electrical Engineering, vol. 526. Springer, Singapore, 2019. DOI: 10.1007/978-981-13-2553-3_2

[27] A. O. Salau, T. K. Yesufu, B. S. Ogundare, "Vehicle Plate Number Localization Using a Modified GrabCut Algorithm," *Journal of King Saud University - Computer and Information Sciences*, 2019. DOI: https://doi.org/10.1016/j.jksuci.2019.01.011