International Journal of Computer (IJC)

ISSN 2307-4531

http://gssrr.org/index.php?journal=InternationalJournalOfComputer&page=index

ISSN: 2307-4523

NEURAL NETWORK MODELS: A Basic Statistical Tool for Data analysis and Forecasting

Maitanmi Olusola [a]*, Malasowe Bridget[b] , Emuobonuvie Erhinyoja Andy[c], Adekunle Yinka[d], Gregory Onwodi[e]

[a]Computer Science Department, Babcock University, P.M.B 21244, Ikeja Lagos, Nigeria

[b]Department of Computer Science, College of Education, Agbor, Delta State, Nigeria

[c]Department of Computer Science, College of Education, Agbor, Delta State, Nigeria

[d]Computer Science Department, Babcock University, P.M.B 21244, Ikeja Lagos, Nigeria

[d]School of Science & Technology, National Open University of Nigeria.
[a]maitanmi@yaoo.com
[b]malasowebrigit@yahoo.com
[c]aemuobonuvie@gmail.com
[d]adekunleya@gmail.com
[e]gregonwodi@hotmail.com

Abstract

Neural network model is a model of brain's cognitive process. Neural network originated as a model of how the brain works and has its beginnings in psychology. Today, neural network models are used to solve numerous problems associated with forecasting, pattern recognition, classification, manufacturing, medical diagnostic, signal processing, system control, checking, and modelling among others. This paper reveals the processes involved in using a neural network model in forecasting future events by effectively and efficiently utilizing extracted rules from trained neural networks and domain knowledge of the applied area. The structural method of operations and rules of extraction of neural networks were also discussed.

Keywords: Neural network, brain cognitive, modelling.

1. Introduction

Forecasting future events are very important to our daily life. Much effort has been made to research on prediction. Statistics and Neural Networks offer the best fit to data and are a satisfactory solution for such problems. Neural Network models are computer software that tries to emulate biological neural networks. A neural network may act as a learning system made up of simple units configured in a highly interconnected network. It is used to solve problems that are hard to find a well formulated algorithmic solution. For instance, we want to find the underlying rules or structures from existing large amount of data or examples.

----
* Corresponding author. Tel.: +2348068856195
E-mail address: maitanmi@yahoo.com

The major difference between neural networks with traditional computing is that neural networks are based on the parallel architecture of animal brains while traditional computing is based on serial and centralized architecture. Hence, neural networks are good candidates for solving prediction problems [4] and [4]. The paper is structured as follows apart from the introduction; structure, methods and operations of neural networks, learning process of neural network, comparison of neural network to conventional computing, extracted rules or knowledge from trained neural networks and application of neural network in forecasting.

1.1 Structure, Methods and Operations of Neural Networks

The human brain is a highly complicated system which is capable of solving complex problems. The brain consists of many different elements, but one of its most important building blocks is the neuron, of which it contains approximately 10^{11} [8]. These neurons are connected by around 10^{15} connections, creating a huge neural network. Neurons send impulses to each other through the connections and these impulses make the brain work.

The artificial analogue of the biological neuron is shown in figure 1 which is referred to as a Processing Element (PE) as shown in figure 2. An artificial neural network consists of small numbers of PE's (tens to thousands). A PE has many input paths (dendrites, $X_{0 \ldots n}$) from adjoining PE's outputs or axons. Input signals are combined usually by simple summation and are passed to the following PE through the axon. Figure 3 shows the basic structure of artificial neural networks. PE's are usually organised into groups. Generally there are three types of layers; the inputs layer collects information presented from the surroundings, the output layer generates a response to a given output while the layer between input and output is called the hidden layer. PE's in any one layer are joined with all PE's in the layer above.

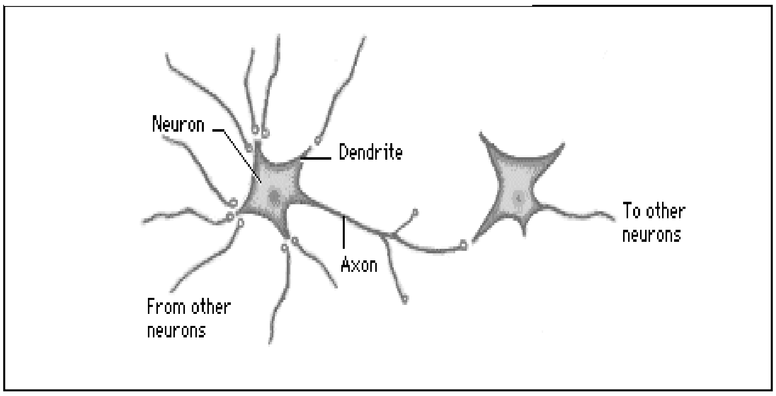

Figure 1 The biological neuron. Source: [6]

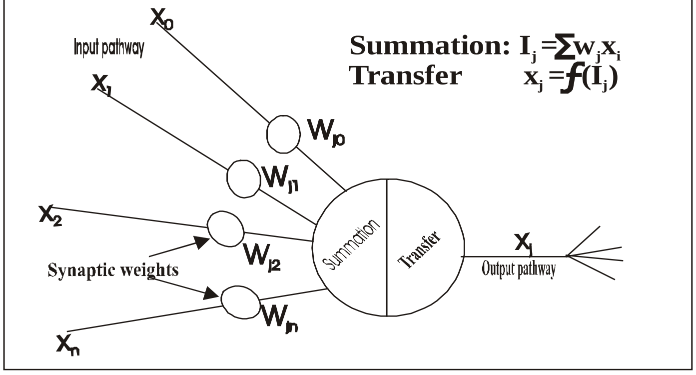

Figure 2. Element Structure and Function. Source: [6]

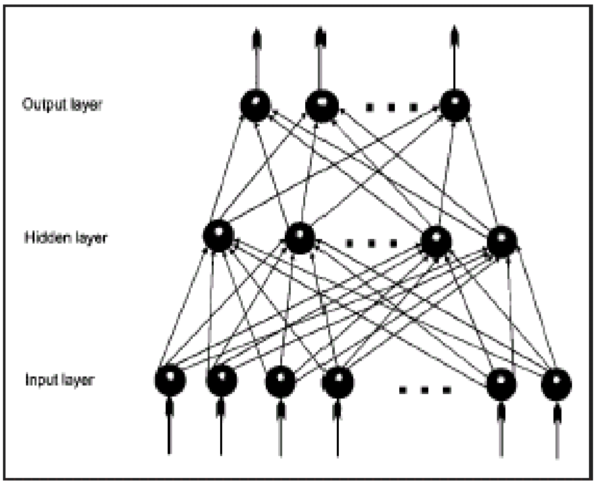

Figure 3. Neural network structure or layers. Source: [6]

Two methods are commonly used in neural network applications; these are neocognitron and back propagation. Neocognitron is a hierarchical network made up of many layers and its organization is like that of the visual cortex. Neocognitron is a method of pattern recognition. It has the capability to recognize shapes and sizes of characters even if it involves noise and distortion. Back propagation allows the training of multi-layer networks, it is a powerful and practical tool for solving problems that would be quite difficult using conventional computer science techniques and these problems range from image processing, speech recognition, character recognition, optimization and forecasting [2].

The operation of a typical back propagation network can be derived as follows:

1. After presenting signals to the input layer, information propagates through the network to the output layer (forward propagation). During this time input and output state for each PE will be set.

   $$x_j^{[s]} = f(I_j^{[s]}) = f\left[\sum_i \left(w_{ij}^{[s]} * x_i^{[s-1]} \right) \right] \quad \text{---------- equation 1}$$

   Where

   $x_j^{[s]}$ denotes the current output state of the j^th^ PE in the current [s] layer,

   $I_j^{[s]}$ denotes the weighted sum of the input of the j^{th} PE in the current [s] layer,

   f is conventionally the sigmoid function

   $w_{ij}^{[s]}$ denotes the connection weight between the i^{th} PE in the current [s] layer and the j^{th} PE in the previous [s-1] layer.

2. Global error is generated based on the summed difference required and calculated output values of each PE in the output layer.

   $$E_{glob} = 0.5 * \sum_{ki} \left((r_k - O_k)^2 \right) \quad \text{-------------- equation 2}$$

   Where E_{glob} means the global error (r_k - o_k) denotes the difference of required and calculated output values.

   Scaled local error for each PE in the output layer is calculated according to the following formula:

   $$e_k^{(o)} = x_k^{[o]} * (1.0 - x_k^{[o]}) * (r_k - o_k) \quad \text{--------------- equation 3}$$

   This formula shows that local errors are scaled based on their output activation values.

3. Global error is back propagated through the network to calculate local error values and delta weight for each PE. Delta weight are modified according to the Delta rule that strictly controls the continuous decrease of the synapse strength of those PE's that are mainly responsible for the global error. In this manner, a regular decrease of global error is assured.

$$e_j^{[s]} = x_j^{[s]} * (1.0 - x_j^{[s]}) * \sum_i \left(e_k^{[s+1]} * w_{kj}^{[s+1]}\right) \text{ ------------- equation 4}$$

where

$e_j^{[s]}$ is the scaled local error of the j^{th} PE in the current [s] layer.

$$\Delta w_{ji}^{[s]} = Lcoef * e_j^{[s]} * x_i^{[s-1]} \text{ ------------------- equation 5}$$

Where

$\Delta w_{ji}^{[s]}$ denotes Delta weight of the connections between the current PE ($PE_j^{[s]}$) and the joining PE ($PE_i^{[s-1]}$)

Lcoef denotes the learning coefficient, one of the training parameters.

4. Synapse weights are updated by adding delta weight ($\Delta w_{ji}^{[s]}$) to the current weights. Neural network simulate neurotransmission by changing the strength of interneural connections (synaptic weights). Positive synaptic weights provide amplified neural signal; and stronger effect to the joining PE. No modification in the information flow is modeled by zero (0) weight. Negative weight means inhibition.

2.0 Learning Process Of Neural Network

Neural network is used to capture the general relationship between variables of a system that are difficult to analytically relate. Neural network has been described as brain metaphor of information processing or as a biologically inspired statistical tool [11]. It has the capacity to learn or to be trained about a particular task, its computational capabilities and the ability to formulate abstractions and generalizations. Neural network has an organization similar to that of the human brain and it is a network made up of processing elements called neurons. Neurons get data from the surrounding neurons, perform some computations, pass the result to other neurons. Connections between the neurons have weight associated with them. In neural network, the knowledge is stored in the networks interconnection weight in an implicit manner, learning take place within the system and plays the most important role in the construction of neural network system. Neural network system learns by determining the interconnections weight from a set of a given data [15]. Research according to Learning [5] reports that neural network can be supervised, unsupervised or based on a combined supervised-unsupervised training which are further expansiated below.

In supervised learning, a set of data called a training data set is used to help the network in arriving at the appropriate weight [15]. A teacher teaches the network and gives results of the output corresponding to the input. The inputs as well as side information indicating the correct output are presented to the network. The network is also programmed to know the procedure to be applied in adjusting the weight and thus the network has the means to determine whether or not its output was correct and the means to apply the learning law to adjust its weight in response to the resulting errors [5].

Weights are generally modified on the basis of the error between desired and actual output in an iterative fashion and one of the widely used training algorithms is the "Delta Rule". The network learns the desired output by adjusting its internal connection weight to minimize the discrepancy between the actual outputs of the system and desired outputs [7].

In unsupervised learning, a network operates in self-organized mode. In this self-organized mode, a competitive mechanism is used to select processing elements due to their weights which are modified. The first part of the competition involves competition between the simple elements that lie within the same sample. Winners from these competition compete with other processing elements in their layer that have won similar competition in order to be single processing element on their layer which has the weight changed during training. In self organizes training, the network is given many input characters and no information is given to the network as to what each example corresponds to. There are many layers and each layer corresponds to one character more than the other layers. The importance of unsupervised learning is that the system does not need to know the correct answer in order to solve a problem. The system learns a pattern from repeated exposure to it and is able to recall the learned pattern when it solves a categorization or pattern matching problem [15].

Neural network can also employ a hybrid approach in which learning is based on combined supervised-unsupervised learning. The hybrid approach first uses unsupervised learning to form clusters and the labels are then assigned to the clusters identified and a supervised training follows [5].

2.1 Comparison Of Neural Network To Conventional Computing

In the course of conventional computational problem solving, software designers and programmers use algorithm-based procedures and linear or non-linear functions. Thus, an algorithm clearly describes the whole problem solving procedure and is then transformed into mathematical functions. On the other hand, neural networks are developed through training (experience), rather than being programmed [9]. Making an appropriate network is relatively simple using computerized simulator software. The well defined network is able to self-adapt to the exercise and prepare its own solving algorithm based on the cyclical processing of the training samples. The problem solving algorithm is included in the neural network synaptic weight matrix. Neural network can improve on conventional rule-based system in their flexibility and ability to adapt as shown in table 1.

There are however, several disadvantages of neural networks compared to algorithm based systems. Although easy to produce tuning a neural network for a specific task requires long experience. There are no appropriate data in literatures describing what type of network to create for a given tasks, how to set learning parameters or how long to train. During training, three types of errors can occur: training the network too accurately will result in a loss of flexible problem solving, with a short training process the network will be unfit for proper accuracy and finally, incorrect learning parameters setting result in reaching a dead end (local minima): this means that training for any further length of time can not lead to a better result. Training accuracy depends not only on the training parameter settings and the number of learning iterations, but also on the samples of the training set.

Table 1: Comparison of algorithm based system to neural networks.

	ALGORITHM BASED SYSTEM	**NEURAL NETWORK**
VARIABLES	Pre-processed parameter	Raw bioelectric signals, pre-processed parameters
CLASSIFICATION	Algorithmic, pre-defined differentiating factor	Training set, training process
REPRODUCIBILITY OF DECISION PROCESS	Available	Not available
DECISION PROCESS	Decision borders and pathways predefined by experts	Automatic self made decision based on learning process

3.0 Extracted Rules From Trained Neural Networks

One of the major drawbacks or challenges of neural network models is that these models can not explain what they have done. Extracting rules from trained neural networks is one of the solutions for understanding the networks [10]. When rules or knowledge are extracted from a trained neural network, they are embedded or used in reconstructing a new neural network called Descriptive Neural Network (DNN). A DNN is a neural network embedded with rules that have been discovered from previously trained networks and description of the domain knowledge of the applied area, so that not only predictions can be made but also the reasons for the predictions can be explained [12]. Hence, the architecture of DNN is not only decided by training examples or data but also by hidden rules extracted from trained network and domain knowledge of applied area. The DNN system to traditional neural network is similar to econometrics to regression analysis. It may not be acceptable by practitioners if only regression or neural network models are used as forecasting tools. Econometrics is the technique of using statistical analysis combined with economics theory to analyze economic data. It presents more economic knowledge than a single mathematical formula and thus more popular and acceptable to practitioners. In fact, it is more accurate than regression models. Similarly, DNN are more accurate than traditional neural networks.

4.0 Applications of Neural Network in Forecasting

In the research conducted by [1], neural network models for predicting poultry performance based on their feed composition were developed. The Generalized Feed Forward Neural Network (GFFNN) model was selected as the one which gave the best performance on the test data as compared to other two network models: Multi-layer perceptron neural network (MLPNN) and RBF/GRNN (Radial basis function/Generalized Regression neural network). The GFFNN was able to learn the problem and uses it to predict the final weight of poultry if they were feed with different feed compositions suggested in the production data set. The GFFNN generated various results. The results indicated that supplementing the maize used in feed composition with as much as 35% of sorghum dust, the addition of exogenous enzyme (â – glucanase) concentration of 0.01, and achieving a high feed efficiency value will produce poultry bird whose weight are close to the standard value of 1.8 kilograms in seven (7) weeks.

In the medical field, neural network have been applied in different areas. These include: Intensive care, Gastroenterology, Obstetrics and Gynaecology among others. Research according to [8] reports that a neural network was described that predicts the presence of tumour in an x-ray image. The network was shown many x-ray images containing tumours, and many x-ray images containing healthy tissues. After a period of training with these images, the network can positively predict and identify tumour in x-ray images that is shown to it.

Further research findings by Cook and [3] revealed that methodology was presented that predicts the occurrence of out-of-control process conditions in a composite board manufacturing facility. This method was developed using neural network theory. The neural network, using back-propagation method, was successfully trained to represent the process parameters. The trained neural network was able to successfully predict the state of control of specific manufacturing process parameters with 70% accuracy. Other application examples include: Using neural network in foreign exchange prediction [14] and stock market assessment and predictions [13].

5.0 Conclusion

The structural method of operations, rules of extraction and application of neural networks are presented in this paper. Ever since the beginning of research in artificial intelligence, neural network has shown tremendous potentials in solving complex problems. Neural network methods based on supervised and unsupervised learning are used in numerous applications such as forecasting, pattern recognition and classification just to mention a few. Solution methods based on neural network are elegant and efficient in solving a wide range of forecasting problems that are critical for survival in today's competitive environment.

Reference

[1]. A. S Ajakaiye, A. B Adeyemo, A. O Osofisan, and O.P A Olowu, O.P.A. (2006). "Analysis of Poultry Birds production performance using Artificial Neural Networks". *Asian Journal of Information Technology its applications,* 5(5) .[Online], Avaialable: http://docsdrive.com/pdfs/medwelljournals/ajit/2006/522- 527.pdf [Sept., 12, 2013]

[2]. A. Blum, "Neural Network in C++" - An object oriented framework for building connectionist system, John Wiley and Sons, New York, 1992 pp. 1-10.

[3]. D. F Cook, and R. E Shannon, (1992). "A predictive neural network modeling system for manufacturing process parameters", *International Journal of production research*, vol. 30, No. 7, pp. 1537-1550. [Online] Available: http://www.tandfonline.com/doi/abs/10.1080/00207549208948106?journalCode=tprs20#preview [Aug.12, 2013]

[4]. M. Duhonx, J. A Suykens, B. Demoor, and J. Vandewalk, (2001). "Improved Long-term temperature prediction by chaining of neural networks" *International Journal of neural systems,* 11(1), pp. 1-10.

[5]. S. Kaparthi, and N. C Suresh (1992). "Machine-component cell formation in Group technology: A neural network approach" *International Journal of Production research, vol. 30, No. 6, pp. 1353 – 1367.*[Online], Available: http://www.tandfonline.com/doi/abs/10.1080/00207549208942961#preview [Sept. 1, 2013]

[6]. Papik K et al (1998). "Application of neural networks in medicine - a review", International journal of Diagnostics and Medical Technology, 4(3): 538-546 [Online], Available: http://uran.donetsk.ua/~masters/2006/kita/zbykovsky/library/nninmed.pdf [June, 12, 2013].

[7]. N. Srinivasa, P. Kalle, J. C Ziegert, and S. Smith, (1992). "Neural Network prediction of machine tool error maps" Proceedings of the National science foundation (NSF) Design and Manufacturing systems, pp. 1099 – 1103.

[8]. N. Steffen, (2005). "Neural Networks made simple", Fast neural network library (Fann). pp. 14-15 [Online] Available:www.software20.org. [Aug. 20, 2013]

[9]. L. Threapleton, and C. I Wilson (2003). "Application of Artificial Intelligence for predicting Beer flavours from chemical analysis", Coors Brewers, Technical centre, P.O. box 12, Cross street, Burton – on-Trent, DE141XH, UK. Presented at the 29th European Brewing Congress on May 17 – 22, at Dublin Ireland.

[10]. A. Tickle, R. Andrews, M. Goles, and J. Diederich, (1998). "The truth will come to light: Directions and challenges in extracting the knowledge embedded within trained Artificial neural networks", *IEEE trans Neural Networks, 9(6), pp. 1057-1068.*

[11]. R. Wilson, and R. Sharda, (1992). "Neural Network" OR/MS TODAY, A joint Publication of RSA and TIMS, pp. 36-42.

[12]. J. T Yao, (2003). "Knowledge Based Descriptive Neural Networks". *Proceedings of the 9th International conference on Rough sets fuzzy sets, Data mining and granular computing, chongqing, china, pp. 272-277.*[Online], Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.14.4944&rep=rep1&type=pdf [Aug. 28, 2013]

[13]. J. T Yao, C. L Tan, and H. L Pol, (1999). "Neural Networks for technical Analysis: A study on KLCI ", *International Journal of theoretical and applied finance*, 2(2), pp. 221-241.[Online], Available: *betterea.googlecode.com/files/klci_ijtaf.pdf* [Sept., 29, 2013]

[14]. J. T Yao, Y. L Li, and C. L , Tan, (1997). "Forecasting the exchange rate of CHF vs USD Using Neural Network", *Journal of computational Intelligence in France, 5(2), pp. 7-13* [Online], Available: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=FA8A60EF1831B126114C86335312DCF6?doi=10.1.1.53.5200&rep=rep1&type=pdf [May, 12. 2013]

[15]. F. Zahedi, (1991). "An Introduction to Neural Networks and a comparison with Artificial intelligence and expert systems" *Interfaces 21:2, pp. 25-38.* [Online], Available : http://www.jstor.org/stable/25061463 [Aug. 23, 2013]