EUROPEAN ACADEMIC RESEARCH
Vol. IX, Issue 2/ May 2021

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

A Review of Artificial Neural Networks Learning Algorithms Used for Prediction

ELDA XHUMARI[1]
PhD Candidate
Department of Informatics, Faculty of Natural Sciences
University of Tirana, Albania
ALBA COMO, PhD[2]
Department of Informatics, Faculty of Natural Sciences
University of Tirana, Albania

Abstract

Nowadays, Artificial Neural Networks (ANNs), based on the attempts to reproduce the human nervous system, are used to solve different problems such as data classification, face and pattern recognition, time series prediction, etc. There are already examples of using ANNs, where, in order to solve the above-mentioned problems, it is necessary to specify the training parameters of the neural network, the network architecture, and the learning algorithms in various types of neural networks. This paper provides an overview of the learning algorithms used for prediction problems based on the fact that ANN is a powerful method for calculating a huge number of relationships between input and output parameters and is able to find complex relationships. The method provides reliable predictive results for many problems through learning by example. The ANN-based model is capable of generating results more accurately than traditional forecasting methods, for example, multilinear regression.

[1] Experienced Lecturer with a demonstrated history of working in the computer software industry. Skilled in OS X, Computer Science and Networking. Strong education professional with two Master of Science (M.Sc.), first one focused in Artificial Neural Networks and the second in Econometrics and Quantitative Economics from Voronezh State University; researcher, conducts courses and seminars, serves as academic advisor to students for: Security and administration in Information Systems; Networks; Informatics; Artificial Intelligence; Neural networks. (elda.xhumari@fshn.edu.al)

[2] Conducts desk research, Research activity is focused on E-learning systems, DBMS systems for reporting analytics data, Date Mining, Big Data. Experienced Lecturer at University of Tirana, researcher, conducts courses and seminars, serves as academic advisor to students for: Database Systems and Administration of Database Systems; Introduction to Computer Science; Introduction to C/C++ Language; Data Structure in C/C++ Language; Logical Systems; Introduction to Big Data. (alba.como@fshn.edu.al)

Keywords: Artificial Neural Networks, Algorithms, Prediction

INTRODUCTION:

The problem of prediction was and remains relevant, especially recently, when powerful tools for collecting and processing information have become available. Prediction or forecasting time series is an important scientific and technical problem, as it allows predicting the behavior of various factors in ecological, economic, social, and other fields (Aghbashlo, Hosseinpour & Mujumdar, 2015).

The development of forecasting as science in recent decades has led to the creation of many models and methods, procedures, forecasting techniques, unequal in value. According to estimates of foreign and domestic experts in forecasting, there are already over a hundred forecasting methods, in connection with which there is the task of choosing methods that would give adequate predictions for the studied processes or systems. Rigorous statistical assumptions about the properties of time series often limit the capabilities of classical forecasting methods. The use of ANNs in this problem is due to the presence of complex patterns in most time series that are not detected by known linear methods (Goncalves et al., 2013).

Neural network methods of information processing began to be used several decades ago. Over time, interest in neural network technologies either waned or revived again. Such inconsistency is related to the practical results of ongoing research. Today, the capabilities of neural network technologies are used in many fields of science, from medicine and astronomy to computer science and economics (Goncalves et al., 2013). The ability of a neural network for versatile information processing follows from its ability to generalize and highlight hidden dependencies between input and output data. The great advantage of neural networks is that they are capable of learning and generalizing accumulated knowledge.

Classification of artificial neural networks:

Artificial neural networks differ in their architecture: the structure of connections between neurons, the number of layers, neuron activation functions, training algorithms. From this point of view, among the wellknown ANNs one can distinguish static, dynamic networks and fuzzy - structures; single-layer and multi-layer networks (Krenek et al., 2014). In general, according to the structure of connections, ANNs can be grouped into two classes: feedforward networks - without feedbacks in the structure and recurrent networks - with feedbacks. In the first class, the most famous and most commonly used are multilayer neural networks, where artificial neurons are arranged in layers. The connection between the layers is unidirectional and, in the general case, the output of each neuron is connected with all the inputs of the neurons of the next layer. Such networks are static because they are having neither feedbacks nor dynamic elements in their structure, and the output depends on a given set at the input and does not depend on the previous states of the network. Second-class networks are dynamic because due to feedbacks, the state of the network at each moment of time depends on the previous state.

Each neural network includes a first layer of neurons called the input layer. This layer does not perform any transformations and calculations, its task is different: to receive and distribute input signals to the rest of the neurons. And this layer is the only one that is common for all types of neural networks, and the further structure is the criterion for division:

- One-layer neural network - a structure of the interaction of neurons, in which signals from the input layer are immediately directed to the output layer, which, not only converts the signal but also immediately produces a response. As already mentioned, the 1st input layer only receives and distributes signals, and the necessary calculations take place already in the second layer. Input neurons are connected to the main layer using synapses with different weights to ensure the quality of the connections (Goncalves et al., 2013).
- Multilayer neural network - in addition to the output and input layers, there are several more hidden intermediate layers. The number of these layers depends on the complexity of the neural network. It is more reminiscent of the structure of a biological neural network. Such types were developed quite recently, before that all processes were implemented using single-layer neural networks. The corresponding solutions have great capabilities when compared with single-layer ones because in the process of data processing, each intermediate layer is an intermediate stage

at which information processing and distribution is carried out (Goncalves et al., 2013).

Figure 1 shows a classification of neural networks according to different attributes.

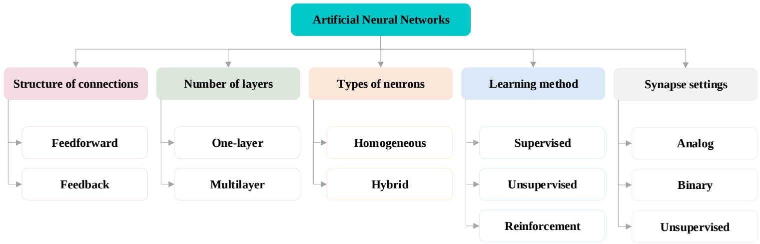

Figure 1. Artificial Neural Networks classification according to different criteria

How ANN works?

The simplest network consists of a group of neurons forming a layer, as shown on the right side in Figure 2. The top circles on the left serve only to distribute the input signals, they do not perform any computation and therefore will not count as a layer. For this reason, they are marked with circles to distinguish them from computing neurons, indicated by squares. Each element of the set of inputs X is connected with a separate weight to each artificial neuron and each neuron produces a weighted sum of inputs to the network. In artificial and biological networks, many connections may be missing, all connections are shown for general purposes. There can also be connections between outputs and inputs of elements in a layer (Krenek et al., 2014).

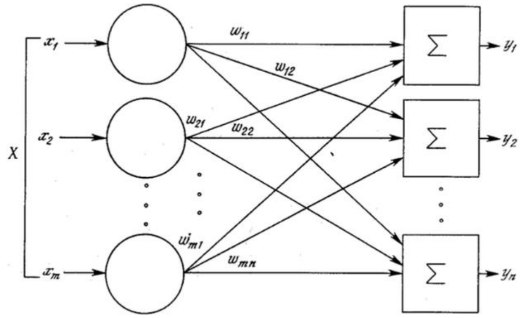

Figure 2. Single-layer linear neural network with weight matrix *W* of size *NxN*

It is convenient to consider the weights as elements of the matrix W. The matrix has m rows and n columns, where m is the number of inputs and n is the number of neurons. For example, is the weight connecting the third input to the second neuron. Thus, the calculation of the output vector N, the components of which are the outputs of the OUT neurons, is reduced to matrix multiplication $N = XW$, where N and X are row vectors (Schmidhuber et al., 2015).

As a rule, the transfer functions of all neurons in the network are fixed, and the weights are the parameters of the network and can change. Some inputs of neurons can be labelled as external inputs of the network, and some outputs can be labelled as external outputs of the network.

When any numbers are fed to the network inputs, a set of numbers is obtained at the network outputs. Thus, the work of a neural network is to transform an input vector into an output vector, and this transformation is set by the network weights.

Learning process:

To get a solution to the problem using an artificial neural network, first of all there is a need to train the network. The process of training a network consists in adjusting the weights of connections between neurons (Figure 3).

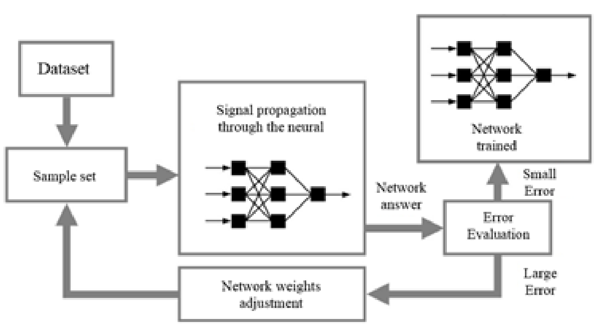

Figure 3. Neural Network Learning Process

In the context of an artificial neural network, the learning process can be viewed as setting up the network architecture and connection weights to efficiently perform a special task (Krenek et al., 2014). Typically, a neural network must adjust the connection weights for the available training sample. Network performance improves as the weights are iteratively adjusted. The property of a network to learn from examples makes them more attractive in comparison with systems that follow certain rules of operation formulated by experts.

There are three paradigms of learning: *(i) supervised, (ii) unsupervised* and *(iii) reinforcement learning* (Happiness, Zhou, Kranthi & Qingtian, 2018). In the first case, the neural network has the correct answers (network outputs) for each input example. The weights are adjusted so that the network produces answers as close as possible to the known correct answers. Unsupervised learning does not require knowing the correct answers for each example of a training set. In this case, the internal data structure or correlations between samples in the data system are revealed, which allows categorizing the samples. In reinforcement learning, some of the weights are determined through supervised learning, while the rest is obtained through self-learning. Reinforcement Learning learns by interconnecting with its surroundings or environment. A Reinforcement Learning manager or agent learns from the significances of its activities, than from being clearly trained and it chooses its actions to base its past information and also by novel choices, which is basically a trial and error learning technique. This is different from classic supervised learning, since precise input/output data sets are not presented. (Happiness, Zhou, Kranthi & Qingtian, 2018). Various reinforcement learning algorithms for this condition uses dynamic programming techniques which is usually demonstrated as a Markov Decision Process (MDP) (Otterlo & Wiering, 2012).

Table 1 presents the various learning algorithms and associated network architectures. The last column lists the type of problems for which each algorithm can be applied. Each learning algorithm is focused on a network of a specific architecture and is designed for a limited class of problems.

Table 1. Neural Network learning algorithms

Category	Training rule	Architecture	Learning algorithm	Type of problem
Supervised	Error Correction	Single-layer and multi-layer perceptron	Perceptron Training Algorithms Back propagation	Image classification Function approximation Prediction
	Boltzmann	Recurrent	Boltzmann Algorithm	Image classification
	Hebb	Multilayer Feedforward	Linear Discriminant Analysis	Data Analysis

	Competition	Competition	Vector quantization	Classification Data compression
		ART	ART Map	Image Classification
Unsupervised	Error Correction	Multilayer Feedforward	Sammon's projection	Classification Data compression
	Hebb	Feedforward or Competition	Principal Component Analysis	Data Analysis Data compression
		Hopfield network	Associative memory training	Associative memory
	Competition	Competition	Vector quantization	Classification Data compression
		Self-organizing Map Kohonen	SOM Kohonen	Data Analysis Data compression
		ART	ART1, ART2	Classification
Reinforcement learning	Error Correction & Competition	Radial Basis Function Network	Radial Basis Function	Image classification Function approximation

Based on the classification presented in Table 1, it turns out that supervised learning algorithms are used to solve prediction problems, which can be classified into two categories:

1. **Classification** - which consists of recognizing specific entities within the database and grouping the data according to a certain attribute. Some of the most common classification algorithms are:
   a. *Linear classifiers* - A popular class of procedures for solving classification problems is based on linear models. This means that they tend to divide the feature space into a set of regions labeled according to the values that the target can take, where the decision boundaries between these regions are linear: lines in 2D, planes in 3D, and hyperplanes with a lot of function (Chen et al., 2007).
   b. *Support Vector Machines* (SVM) can be used to classify and regression problems like vector classification support (SVC) and vector regression support (SVR). It is used for a smaller dataset because it takes too long to process. SVM is based on the idea of finding a hyperplane that best separates elements into different regions (Chen et al., 2007).
   c. *Decision trees*, which are hierarchical tree structures consisting of decision rules of the form "If ... then ...". The rules are automatically generated in the learning process on the training set (Chen et al., 2007).
   d. *K-nearest neighbor* - It assigns objects to the class where most of its k nearest neighbors belong, in a multidimensional feature space. This is one of the simplest algorithms for training classification models. The number k is the number of neighboring objects in the feature space,

which are compared with the classified object (Chen et al., 2007).

e. *Random forest* - The random forest algorithm creates decision trees for samples of data and then obtains a prediction for each of them, and at the end chooses the best solution through a vote. This ensemble method is better than a single decision tree because it reduces overfitting by averaging the result (Chen et al., 2007).

2. **Regression** - used to understand the relationship between independent and dependent variables. Some of the most common regression algorithms are (Liu et al., 2003):
    a. *Linear regression* - Models and analyzes relationships between variables, and to see how these variables together affect a particular outcome.
    b. *Logistical regression* - Models the probabilities of an observation belonging to each of the classes through linear functions ensuring that these probabilities sum to one and remain in the range (0, 1).
    c. *Polynomial regression* – Models the relationship between a dependent and independent variable as n^{th} degree polynomial.

CONCLUSION:

There are numerous artificial neural networks utilized to design and validate machine learning components across diverse fields, such as image classification, function approximation, prediction, data analysis, data compression, etc. (Foody et al., 2013). The discussed learning algorithms are primarily categorized based on paradigm of learning, type of problems, training rules and ANN architecture. Comparison of learning algorithms for ANN for prediction or forecasting problems is summarized in Table 2 below, in accordance with their advantages and disadvantages.

Table 2. Summary of ANN learning algorithms used for prediction

No	Learning algorithm	Advantages	Disadvantages
1	Perceptron Learning Algorithms	ease of implementation; fast learning algorithm	Primitive algorithm, able to solve only the simplest problems
2	Back propagation	ease of implementation; robustness to anomalies and outliers in data.	indefinitely long learning process; the possibility of "paralysis of the network", when, at large values, the weights are hardly corrected and the learning process "freezes"; the vulnerability of the algorithm to falling into the local minima of the error function

REFERENCES:

1. Aghbashlo, M., Hosseinpour, S., & Mujumdar, A. (2015). *Application of Artificial Neural Networks (ANNs) in Drying Technology: A Comprehensive Review.* Drying Technology, 33(12), 1397-1462. DOI: 10.1080/07373937.2015.1036288
2. Goncalves, V., Maria, K., & da Silv, A. (2013). *Applications of Artificial Neural Networks in Chemical Problems. Artificial Neural Networks*. Architectures And Applications, 203-223. DOI: 10.5772/51275
3. Krenek, J., Kuca, K., Krejcar, O., & Maresova, P. (2014). *Artificial Neural Networks Tools for Computerised Data Modeling and Processing*. CINTI2014. Retrieved from https://www.researchgate.net/publication/268995502_Artificial_Neural_Network_Tools_for_Computerised_Data_Modeling_and_Processing
4. Schmidhuber J., (2015). *Deep learning in neural networks: an overview*. Neural Netw 61:85–117
5. Happiness U., Yimin Zh., Kranthi K., & Qingtian W. (2018). *Unsupervised Learning Based On Artificial Neural Network: A Review.* Proceedings of the 2018 IEEE International Conference on Cyborg and Bionic Systems. Retrieved from IEEE Xplore
6. Van Otterlo M., Wiering M. (2012). *Reinforcement learning and Markov Decision Processes.* In: Reinforcement learning. Springer, pp 3–42
7. Chen, J., *A comparison of four data mining models: Bayes, neural network. SVM and decision trees in identifying syndromes in coronary heart disease*, 2007. 4491: p. 2007.
8. Liu, C., et al., *Comparison of neural networks and statistical methods in classification of ecological habitats using FIA data*. Forest Science, 2003. 49(4): p. 619-631
9. Foody, G., *Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes.* International Journal of Remote Sensing, 2004. 25(15): p. 3091-3104.