RESEARCH ARTICLE OPEN ACCESS

Artificial Neural Network : A Brief Overview

Magdi Zakaria, Mabrouka AL-Shebany, Shahenda Sarhan
Sirte University SIRTE, LIBYA

Abstract
Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. Neural networks, have remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. so in this paper we tried to introduce a brief overview of ANN to help researchers in their way throw ANN.

I. Introduction

An Artificial Neural networks are very powerful brain-inspired computational models. Which have been employed in various areas such as computing, medicine, engineering, economics, and many others.An artificial neural network is based on the optimization theory. An Artificial Neural Network is a computational model inspired in the functioning of the human brain. It is composed by a set of artificial neurons (known as processing units) that are interconnected with other neuron these neurons depend on weights of the neural network. As The word network in Neural Network refers to the interconnection between neurons present in various layers of a system. These weights represent the connections between the neurons which determine the impact of one neuron on another[1].

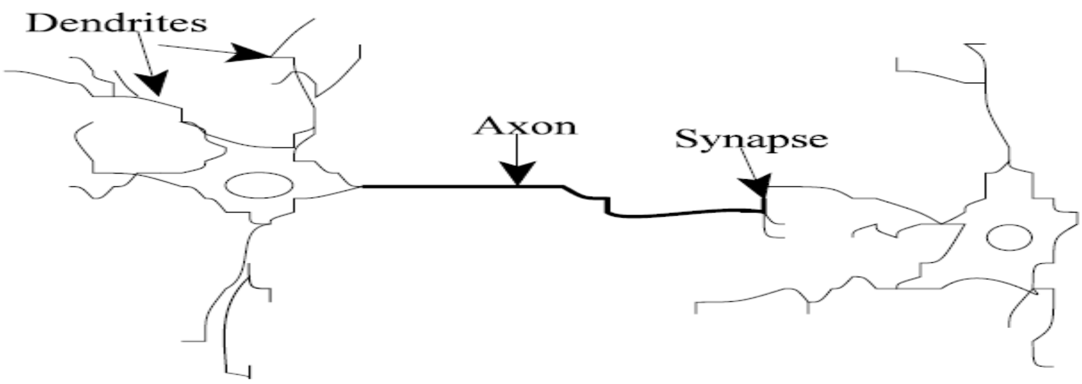

Fig. 1. Natural neuron

The first artificial neuron was firstly proposed in a formal model in 1943 by McCulloch and Pitts. They proved that this model of neuron was able to perform any computable function using a finite number of artificial neurons and synaptic weights adjustable.

The neurons in the input layer receive the data and transfer them to neurons in the first hidden layer through the weighted links. Data are mathematically processed and transferred the result to the neurons in the next layer. The network's output is provided by the neurons in the last layer. The j-th neuron in a hidden layer processes the incoming data (xi) by: (i) calculating the weighted sum and adding a "bias" term (θj) according to:

$$net_j = \sum_{i=1}^{m} x_i * w_{ij} + \theta_j \qquad j = (1, 2, 3, n) \qquad [2]$$

neural networks are capable of adaptation to given data; neural networks are capable of generalization even when the input data set contains noise or missing values (trained network has the capability of correctly filling the value without affecting the prediction); neural networks act as an universal approximation for an arbitrary continuous function with arbitrary precision.

ANN can be defined based on the following three characteristics:

1. The Architecture indicating the number of layers and the no. of nodes in each of the layers.
2. The learning mechanism applied for updating the weights of the connections.
3. The activation functions used in various layers [3].

- **Neural network layers:**

1. Dimensional layer (Single Layer Perceptron) is a function of N real variables of the form

$$f(x_1, \ldots\ldots, x_N) = sgn(\sum_{i=1}^{N} w_i x_i - \theta)$$

Here xi are real variables, (x_1, . . . , x_N) takes values in some domain U ⊂RN, $w_i$are real parameters(weights of the neuron), _ is the threshold of activation of the neuron, the function sgm(x) = 1 for x ≥ 0 and is equal to zero for x < 0.

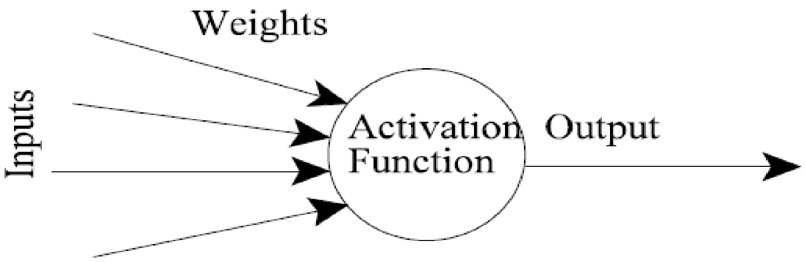

Fig. 2. An artificial neuron

We also consider the smoothed variant of the above neuron for which instead of the function sgm we use the smooth monotonous increasing function sgm which varies from zero to unity. In particular we consider the neuron of the form

$$f(x_1, \ldots\ldots, x_N) = sgn\left(\sum_{i=1}^{N} w_i x_i - \theta\right), sgm(x) = \frac{1}{1+e^{-x}}$$

2. Multi Layer Perceptron (MLP) : Multi-layer consisting of three consecutive layers: an input, a hidden, and an output layer. Every system is basically a three layered system, which are Input layer, Hidden Layer and Output Layer. The input layer has input neurons which transfer data via synapses to the hidden layer, and similarly the hidden layer transfers this data to the output layer via more synapses. The synapses stores values called weights which helps them to manipulate the input and output to various layers.

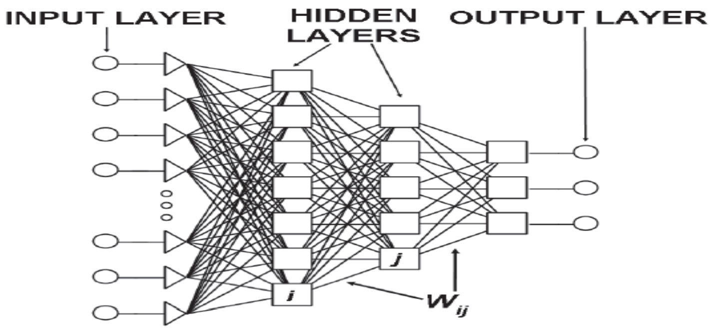

Fig.3. General structure of a neural network with two hidden layer

II. Learning in ANN

There are three major learning paradigms; supervised learning, unsupervised learning and reinforcement learning. Usually they can be employed by any given type of artificial neural network architecture. Each learning paradigm has many training algorithms.

- ***Supervised learning***

Supervised learning is a machine learning technique that sets parameters of an artificial neural network from training data. The task of the learning artificial neural network is to set the value of its parameters for any valid input value after having seen output value. The training data consist of pairs of input and desired output values that are traditionally represented in data vectors. Supervised learning can also be referred as classification, where we have a wide range of classifiers, each with its strengths and weaknesses.

Choosing a suitable classifier (Multilayer perceptron, Support Vector Machines, k-nearest neighbor algorithm, Gaussian mixture model, Gaussian, naive Bayes, decision tree, radial basis function classifiers,…) for a given problem is however still more an art than a science. In order to solve a given problem of supervised learning various steps has to be considered. In the first step we have to determine the type of training examples[6][14]. In the second step we need to gather a training data set that satisfactory describe a given problem. In the third step we need to describe gathered training data set in form understandable to a chosen artificial neural network. In the fourth step we do the learning and after the learning we can test theperformance of learned artificial neural network with the test (validation) data set. Test data set consist of data that has not been introduced to artificial neural network while learning.

2.2 Unsupervised learning

Unsupervised learning is a machine learning technique that sets parameters of an artificial neural network based on given data and a cost function which is to be minimized. Cost function can be any function and it is determined by the task formulation. Unsupervised learning is mostly used in applications that fall within the domain of estimation problems such as statistical modeling, compression, filtering, blind source separation and clustering[14].In unsupervised learning we seek to determine how the data is organized. It differs from supervised learning and reinforcement learning in that the artificial neural network is given only unlabeled examples.

One common form of unsupervised learning is clustering where we try to categorize data in different clusters by their similarity. Among above described artificial neural network models, the Self-organizing maps are the ones that the most commonly use unsupervised learning algorithms.

2.3 Reinforcement learning

Reinforcement learning is a machine learning technique that sets parameters of an artificial neural network, where data is usually not given, but generated by interactions with the environment. Reinforcement learning is concerned with how an artificial neural network ought to take actions in an environment so as to maximize some notion of long-term reward. Reinforcement learning is frequently used as a part of artificial neural network's overall learning algorithm.

After return function that needs to be maximized is defined, reinforcement learning uses several algorithms to find the policy which produces the maximum return. Naive brute force algorithm in first step calculates return function for each possible policy and chooses the policy with the largest return. Obvious weakness of this algorithm is in case of extremely large or even infinite number of possible policies[1][14]. This weakness can be overcome by *value function approaches* or *direct policy estimation*. Value function approaches attempt to find a policy that maximizes the return by maintaining a set of estimates of expected returns for one policy; usually either the current or the optimal estimates. These methods converge to the correct estimates for a fixed policy and can also be used to find the optimal policy.

Similar as value function approaches the direct policy estimation can also find the optimal policy. It can find it by searching it directly in policy space what greatly increases the computational cost. Reinforcement learning is particularly suited to problems which include a long-term versus short-term reward trade-off. It has been applied successfully to various problems, including robot control, telecommunications, and games such as chess and other sequential decision
making tasks.

III. Neural Network Algorithms:

- ***Radial basis functions network (RBF)***

A radial basis function network is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters.

- ***Auto correlation function***

Autocorrelation is the cross-correlation of a signal with itself. It is the similarity between observations as a function of the time interval between them. It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals.

- ***Self-organizing map (SOM) neural network***

A self-organizing map (SOM) is a type of artificial neural network (ANN) that is trained using unsupervised learning to produce a low-dimensional (typically two-dimensional), discretized representation of the input space of the training samples, and called a map. Self-organizing maps are different from other artificial neural networks in the sense that they use a neighborhood function to preserve the topological properties of the input space. The ideas of the SOFM combined with the elastic net algorithm to solve Euclidean problems like the travelling salesman problem. A modified SOFM has been used to solve broad classes of optimization problems by freeing the technique from the Euclidean plane [5].

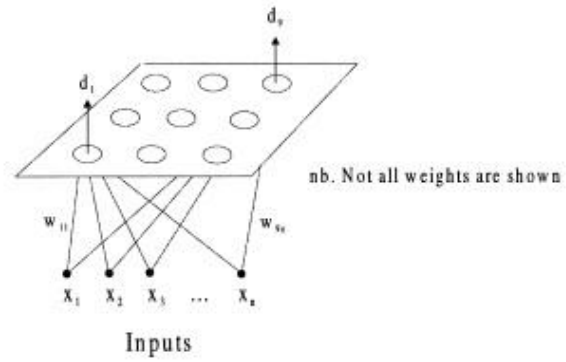

Fig. 4. Architecture of a SOFM with nine neurons

- ***Back propagation neural***

Back propagation ANNs represented by nonlinear networks. The back propagation (BP) algorithmused for training multilayer networks by means of error propagation via variational calculations. It minimizes the sum of squared approximation errors using a gradient descent technique.When noisy training data are present, the learned function can oscillate abruptly between data points. This is clearly undesirable for function approximation from noisy data [6].

- ***<u>Hop-field networks</u>***

Hop-field networks used to solve optimization problems which are familiar to theoperations researcher. although there is no layerstructure to the architecture, and the weights are constants and symmetric. Hop-field networks are a fullyinterconnected system of N neurons. The weights of the network store information about the memories or stablestates of the network. Each neuron has a state xi which is bounded between 0 and 1. Neurons areupdated according to a differential equation, and an energy function is minimized over time because of the stable states of the network.

Hop-field and Tank showed that the weights of a Hop-field network can be chosen so the process of neurons minimizes the Hop-field energyfunction and the optimization problem.

Each neuron i updates itself according to the differential equation

$$\frac{\mathrm{dnet_i}}{\mathrm{dt}} = -\frac{\mathrm{net_i}}{\tau} + \sum_{j=1}^{N} w_{ij} x_j + I_i$$

x_i=f(net_i) Where f(.) is a sigmoidal output function bounded by 0 and 1 and q is a constant [5].

IV. Neural Network Applications:

- *<u>Classification</u>* : such as classification of brain tumors based on many sources of information. KailashD.Kharat&PradyumnaP.Kulkarni&M.B. Nagori proposed method that can do classification of brain tumors by the analysis of Magnetic Resonance (MR) images and Magnetic Resonance Spectroscopy (MRS) data of the patients with benign and malignant tumors to determine the type of tumors [7].

- *<u>Image Indexing And Retrieval</u>*. SooBeom Park, Jae Won Lee, Sang Kyoon Kim proposed a content-based image classification method that reflect the shape of an object based on pattern of the texure feature[8].

- *<u>Pattern Recognition</u>* : hierarchical neural networks improves supervised pattern classification as it used in online back propagation, improved records on MNIST, Latin letters, Chinese characters, and traffic signs. Dan Ciresan ,Ueli Meier, J¨urgenSchmidhuber proposed a system to get data sets ranging from handwritten digits (MNIST), handwritten characters to 3D toys (NORB) and faces[9].

- *<u>Space Exploration Robotic Vehicle And Exploring The Land Of A Planet</u>*. It has the capability to travel across the surface of a landscape and other cosmic bodies. Artificial neural networks have many advantages in space applications due to its:
- Generality
- Performance
- Adaptability
- Low energy consumption

- Robustness & Fault Tolerance: Youssef Bassil proposes a path-planning solution for autonomous robotic planetary rover systems based on artificial neural network (ANN). The proposed ANN uses a mix of activation functions including Sigmoid for the hidden neurons and linear for the output neurons [10]. Where Madhusmita Swain, Sanjit Kumar Dash, Sweta Dash and AyeskantaMohapatra Proposed The Multi-Layer Feed Forward Neural network which is able to classify the three different types of IRIS plants of 150 instances with just few errors for the other one[3].
- *Wavelet Feature Extraction* : proposed for image classification problem. jay Kumar Singh, Shamik Tiwari, V.P. Shukla proposed feature extraction and classification of multiclass images by using Haar wavelet transform and back propagation neural network [11].

- *Multi-class object recognition is a critical capability for an intelligence robot to perceive its environment.* YuhuaZheng and Yan Meng Proposed a model combined a number of modular neural networks to recognize multiple classes of objects for a robotic system. The population of the modular neural networks depends on the class number of the objects to be recognized and each modular network only focuses on learning one object class. For each modular neural network, both the bottom-up (sensory-driven) and top-down (expectation-driven) pathways are attached together, and a supervised learning algorithm is applied to update corresponding weights of both pathways. Also there are two different training strategies are evaluated: positive-only training and positive-and-negative training [12]

- *Pattern Classification* WawanSetiawan and Wiweka proposed a classifier model with neural network approach based on the method used Expectations Maximum (EM)[13].

V. Conclusion

The computing world has a lot to gain from neural networks. Their ability to learn by example makes them very flexible and powerful. Furthermore there is no need to devise an algorithm in order to perform a specific task; i.e. there is no need to understand the internal mechanisms of that task. They are also very well suited for real time systems because of their fast response and computational times which are due to their parallel architecture.

References

[1] 1S.V. Kozyrev," Classification by ensembles of neural networks", Steklov Mathematical Institute, February 21, 2012.

[2] Filippo Amato, Alberto López, Eladia María Peña-Méndez, Petr Vaňhara, Aleš Hampl, Josef Havel, "Artificial neural networks in medical diagnosis", Journal of APPLIED BIOMEDICINE, J Appl Biomed. **11**: 47–58, 2013

[3] Madhusmita Swain, Sanjit Kumar Dash, Sweta Dash and Ayeskanta Mohapatra, "AN APPROACH FOR IRIS PLANT CLASSIFICATION USING NEURAL NETWORK", International Journal on Soft Computing (IJSC) Vol.3, No.1, February 2012

[4] Jose´ L. Herrera-Aguilar, Herna´n Larralde, Maximino Aldana," Pattern Recognition in Neural Networks with Competing Dynamics: Coexistence of Fixed-Point and Cyclic Attractors", PLOS ONE, 1 August 2012, Volume 7, Issue 8, e42348

[5] Kate A. Smith, Jatinder N.D. Gupta,"Neural networks in business: techniques and applications for the operations researcher", Computers & Operations Research 27 (2000) 1023}1044

[6] P.SIBI, S.ALLWYN JONES, P.SIDDARTH," ANALYSIS OF DIFFERENT ACTIVATION FUNCTIONS USING BACK PROPAGATION NEURAL NETWORKS", Journal of Theoretical and Applied Information Technology 31st January 2013. Vol. 47 No.3

[7] Kailash D. Kharat & Pradyumna P. Kulkarni & M. B. Nagori ,"Brain Tumor Classification Using Neural Network Based Methods", International Journal of Computer Science and Informatics ISSN (PRINT): 2231 –5292, Vol-1, Iss-4, 2012.

[8] SooBeom Park, Jae Won Lee, Sang Kyoon Kim, "Content-based image classification using a neural network", Pattern Recognition Letters 25 (2004) 287–300

[9] Dan Ciresan, Ueli Meier, J¨urgen Schmidhuber, "Multi-column Deep Neural Networks for Image Classification" , Technical Report No. IDSIA-04-12 February 2012.

[10] Youssef Bassil, " Neural Network Model for Path-Planning Of Robotic Rover Systems", *International Journal of Science and Technology (IJST)*, E-ISSN: 2224-3577, Vol. 2, No. 2, February, 2012

[11] jay Kumar Singh,Shamik Tiwari, V.P. Shukla, " Wavelet based Multi Class image

classification using Neural Network" , *International Journal of Computer Applications (0975 – 8887) Volume 37– No.4, January 2012*

[12] YuhuaZheng and Yan Meng, " Modular Neural Networks for Multi-Class Object Recognition", 2011 IEEE International Conference on Robotics and Automation Shanghai International Conference Center May 9-13, 2011, Shanghai, China

[13] WawanSetiawan and Wiweka, "Combining of Image Classification With Probabilistic Neural Network (PNN) Approaches Based on Expectation Maximum (EM)", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012

[14] WWW. Wikipedia.or