ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

Simplified Approach of ANN: Strengths and Weakness

Swati G.Anantwar, Rajeshri R. Shelke

Abstract: ***The development of Neural Networks has been so rapid that they are now referred as the sixth generation of computing. While the main strength of Neural Networks is embedded in its non-linearity and data-driven aspects, its main shortcoming relates to the lack of explanation power in the trained networks due to the complex structure of the networks. This paper explains the distinct mechanisms embodied in Neural Networks, its strengths, weaknesses and applications. This paper is for those readers with little or no knowledge of ANNs to help them understand the other articles in this issue of Computer. We discuss The motivations behind the development of ANN" s , describe the basic Biological neuron and the artificial computational model, outline network architectures and learning processes, and present some of the most commonly used ANN models. We conclude with Neural Networks along with explanations on its different components. Neural Networks have gained so much ground that they are now termed as the sixth generation of computing.***

Keywords**: Neural Networks, Artificial Intelligence, Perceptron, Hidden Layers, Neurodynamics, Weights, Node, Over Fitting,**

I.INTRODUCTION

Considered as a subset of Artificial Intelligence, NN basically constitutes a computer program designed to learn in a manner similar to the human brain. Axons which have an electrical signal and found only on output cells, terminate at synapses that connect it to the dendrite of another neuron. NN, throughout the Abilities of human brain.. However, as NN displayed overall part of the analysis, pertains to ANN and not biological NN. NN take after the human brain on the following two grounds, first, they acquire knowledge via the network through a learning process and secondly, interneuron connection strengths are used to store the acquired knowledge. In a nutshell, NN tries to sieve out patterns present in the past data and to extrapolate them into the future. Prior to gaining an insight of NN, it is important to understand the basic components of NN.

II. NEURAL NETWORK SIMPLIFIED

A. Biological neural network: A neuron (or nerve cell) is a special biological cell that processes information (see Figure 1). It is composed of a cell body, or soma, and two types of out-reaching tree-like branches: the axon and the dendrites. The cell body has a nucleus that contains information about hereditary traits and plasma that holds the molecular equipment for producing material needed by the neuron. A neuron receives signals (impulses) from other neurons through its dendrites (receivers) and transmits signals generated by its cell body along the Axon (transmitter), which eventually branches into strands and sub strands. At the terminals of these strands are the synapses. A synapse is an elementary structure and functional unit between two neurons (an axon strand of one neuron and a dendrite of another), When the impulse reaches the synapse's terminal, certain chemicals called neurotransmitters are released. The neurotransmitters diffuse across the synaptic gap, to enhance or inhibit, depending on the type of the synapse, the receptor neuron's own tendency to emit electrical impulses. The synapse's effectiveness can be adjusted by the signals passing through it so that the synapses can learn from the activities in which they participate. This dependence on history acts as a memory, which is possibly responsible for human memory.

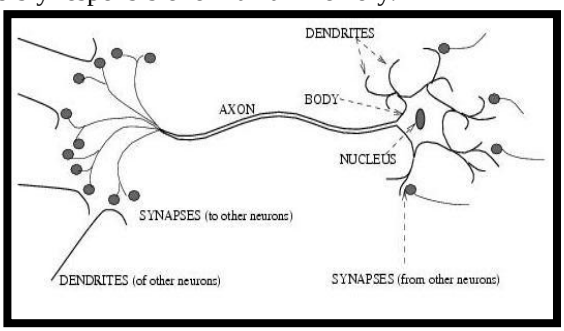

Fig1: Biological Neuron

B. Artificial neuron: An artificial neuron is a mathematical function conceived as a crude model, or abstraction of biological neurons. Artificial neurons are the constitutive units in an artificial neural network. Depending on the specific model used, it can receive different names, such as semi-linear unit, Nv neuron, binary neuron, linear threshold function or McCulloch–Pitts (MCP) neuron . The artificial neuron receives one or more inputs (representing the one or more dendrites) and sums them to produce an output (representing a biological neuron's axon). Usually the sums of each node are weighted, and the sum is passed through a non-linear function known as an activation function or transfer function. The transfer functions usually have a sigmoid shape, but they may also take the form of other non-linear functions, piecewise linear functions, or step functions.

C. Network Architecture:

The simplest form of NN consists of only two layers, the input and output layer (no hidden layer is present). This is sometimes referred to as the skip layer, which basically constitutes a conventional linear regression modeling in a NN design whereby the input layer is directly connected to the output Layer, hence bypassing the hidden layer. Like any other network, this simplest

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

form of NN relies on Weight as the connection between an input and the output; the weight representing the relative Significance of a specific input in the computation of the output. It is important to bear in mind that the Output generated will heavily depend on the type of activation function used. However, based on the Fact that the hidden layer confers strong learning ability to the NN, in practical applications, a three and above three NN architecture is used.

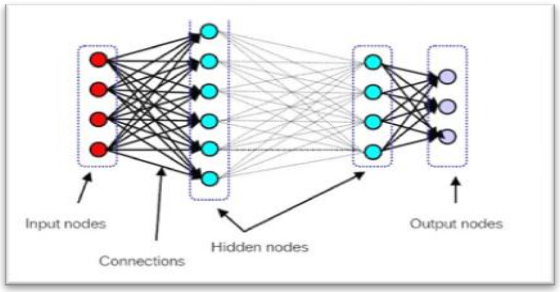

Fig 2: Artificial Neural Network

This is shown in figure 2. It is vital to distinguish between two classes of weights in NN; first there are those weights that aim to connect the inputs to the hidden layer and then those weights that connect the hidden layer to the output layer. In a parallel manner, there are two classes of activation functions, one found in the hidden layer and one in the output layer An infinite number of ways prevail as to the construction of a NN; neurodynamics (basically Spells out the properties of an individual neuron such as its transfer function and how the inputs are Combined) and architecture (defines the structure of NN including the number of neurons in each layer And the number of types of interconnections) are two terms used to describe the way in which a NN is Organized. Any network designer must factor in the following elements when building up a network:

- Best starting values (weight initialization)
- Number of hidden layers
- Number of neurons in each hidden layer
- Number of input variables or combination of input variables
- Learning rate
- Momentum rate
- Training time or amount of training(i.e., the number of iterations to employ)
- Type of activation function to use in the hidden and output layers

Ideally, a NN for a particular task has to be optimized over the entire parameter space of the Learning rate, momentum rate, number of hidden layers and nodes, combination of input variables and Activation functions. Attaining such an objective is computationally burdensome. There is widespread Consensus in the empirical literature that the final artifact of NN model rests purely on a process of trial And error as there is poor theoretical guidance in creating a NN network. In that respect, the design of a network is considered an art rather than a science.

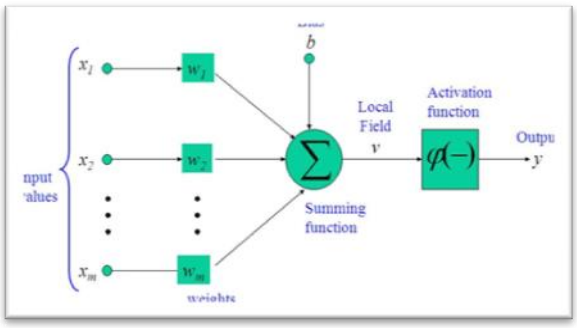

Fig 3: Working of ANN

***1. Weight Initialization*:** Each connection has an associated parameter indicating the strength of the connection, the so-called weight. By changing the weights in a specific way, the network can learn patterns present in the input layer to target values or the output layer.

***2. Number of Hidden Layers*:** The significance of the hidden layer is that; from it, one can infer the number of layers in the NN architecture. For instance, a one-hidden layer signifies a three-layer NN model; a two-hidden layer signifies a four-layer NN model and so forth. Though hidden layers don't present any real concept or that they have no interpretation, yet, they were extremely powerful in sieving distinct pattern structures present in the data. Raising the number of hidden layers not only scales up the computation time but also the probability of Over fitting which then paves the way towards poor out of sample forecasting performance. The usual recommendation employed is either to begin with one or at most two hidden layers with additional layers not recommended since it has been found that NN with 4-hidden layers do not improve results.

3. Number of Units or Neurons in the Hidden Layer: One of the greatest issues in designing NN model lies in deciding on the number of hidden units in the Hidden layer as this is important to ensure good generalization. The fewer the neurons in a network, the fewer the number of operations required and this is a less time-consuming task to implement. In case the number of hidden neurons is too few, the model will not be flexible enough to model the data well. Conversely, in case of too many hidden neurons, the model will over fit the data.

4. Number of Inputs: The inputs should have a temporal structure and should not be too numerous. In case of n independent variables, then, there should be n number of input nodes for the proposed model. Usually, the practice is to go for simple models which have few input variables. The reason is that, as the number of model inputs scales up, the degrees of freedom of the governing equation also increases. While equations with high degrees of freedom have the capability to model the training data effectively, they fail miserably when given test data. This is because models with fewer degrees of freedom do not try to trace the data's random scattering but only follow the general trend.

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

5. Learning Rate: NN models can be classified by whether the learning or iterative adjustment is supervised or Unsupervised.

Supervised Learning:

- The network is provided with a correct answer (output) for every input pattern
- Weights are determined to allow the network to produce answers as close as possible to the known correct answers
- The back-propagation algorithm belongs into this category.

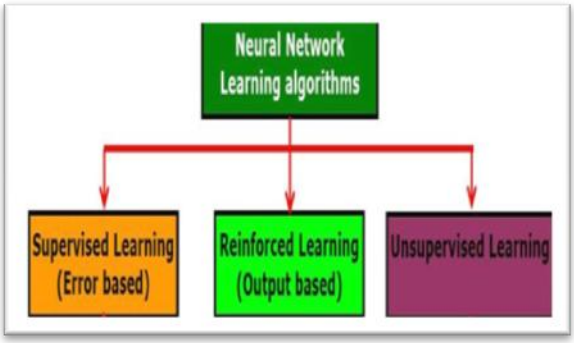

Fig 4: Learning in ANN

Unsupervised Learning:

- Does not require a correct answer associated with each input pattern in the training set
- Explores the underlying structure in the data, or correlations between patterns in the data, and organizes patterns into categories from these correlations
- The Kohonen algorithm belongs into this category

Reinforced Learning:

- Combines supervised and unsupervised learning
- Part of the weights are determined through supervised learning and the others are obtained through a unsupervised learning

6. Momentum Rate: The momentum rate is incorporated to modify the back propagation algorithm. The objective behind the use of the momentum rate is basically to reach the minimum faster because if such a momentum term is excluded, it takes a long time before attaining the minimum. Technically speaking, the momentum term determines how past weight changes impact on current weight changes so that it, in effect, restrains side to side oscillations by filtering out high frequency variations. The momentum stimulates weight changes to continue in the same direction making the current weight change depend not only on the previous weight change but also on the current error.

7. Training **:** Training pertains to the process during which data is input into the network along with their corresponding output values so that the network can adjust the weights in such a way that it can generate the given input and output vectors with as low error as possible. Training will be done until convergence manifests at the point where the test set error start to rise to the effect that the network weights are restored at the iteration cycle where the test set error was minimum. Training should only be ceased until there is no improvement in the error function based on a sensible number of randomly selected starting weights. Convergence constitutes the point at which the network does not improve. The purpose of convergence training is to end up with a global minimum. Training is affected by many Parameters such as selection of the learning rate, momentum values and the back propagation Algorithm.

8. Activation Function (squashing function):

The selection of the squashing or activation function bears an important effect on the NN results. The Activation function constitutes a mathematical transformation of the summarized weighted input units to generate the output of the neuron. Usually, studies use sigmoid activation function for the hidden Layer and linear or sigmoid one for the output layer. Technically speaking an activation function comprises of two parts; a combination function that factors in all the input units into a single value (Weighted sum of inputs) and a transfer function which applies a nonlinear transformation to these Summarized units to trigger an output unit. Usually, the data pattern determines the forms of the transfer functions. Most studies resort towards the sigmoid transfer functions. The benefits of logistic or sigmoid function is that they are continuous functions that monotonically rise or fall, saturate towards the minimum and maximum values, approximate the step function very well and are differentiable on the whole domain and can thereby dramatically reduce the computation burden for training It is important to note that NN where the hidden neurons have sigmoidal activation function and the output neurons the sigmoidal or identity function are called Multi Layer Perceptrons (MLP)

D. Types of ANN:

An Artificial Neural Network can be represented using a directed graph G. There are many classes of Neural Net Works. And they are classified accord nary to their leaning machinery. Mainly they are classified into three fundamentally different classes.

A. Feed Forward Network
   1. Single Layer Artificial Neural Network
   2. Multi Layer Artificial Neural Network

B. Recurrent Network

C. Competitive Network

E. Choice of Network Topology:

Too simple network topologies which have too few connections between their elements have little capacity to store knowledge about the regularities in the training data. On the other side, more complex network topologies will tend to have more weights that are adjusted in the training process of the network. The rationale for the overwhelming use of multilayer perceptron (MLP) is due to the fact that such network architecture is akin to multivariate non-linear regression model. The most common neural network model is the MLP, also known

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

as supervised network by virtue of its need for a targeted output to be able to learn. The aim of such a network is geared towards building up a NN model that properly maps the input to the output.

F. Data Preprocessing:

As a matter of fact, the reduction of dimensionality of the training data is vital for a proper functioning of NN. Another technique is to use ratios. Sampling or filtering the data refers to removing observations from the training and testing sets to generate a more uniform distribution. Benefit of filtering is a fall in the number of training facts which enables testing of more input variables, random starting weights or hidden neurons rather than training large data sets. In practice, data preprocessing involves much trial and error.

G. Strengths of ANN:

The following are deemed as the strengths of NN:

A. The greatest power of Neural Networks is that it is endowed with a finite number of hidden units, can yet approximate any continuous function to any desired degree of accuracy. This has been commonly referred to as the property of universal approximator.
B. No prior knowledge of the data generating process is needed for implementing NN
C. Problem of model misspecification does not occur
D. In case of NN since no specifications are used as the network merely learns the hidden relationship in the data.

H. Drawbacks of ANN:

Accessing the other side of the coin, the following drawbacks were noted for NN:

A. The addition of too many hidden units incites the problem of over fitting the data; meaning that the network learns too well in the training data session but generates inferior results in case of out of sample session.
B. The construction of the NN model can be a time-consuming process since building up the NN architecture is synonymous to a strenuous activity involving trial and error.
C. NN should never be viewed as a panacea.

9. Applications of NN in Distinct Fields:

Neural networks can be applied to almost any problem where you have a) historical data and b) a need to create a model for that data. Neural networks have been successfully applied to broad spectrum of data-intensive applications, such as:

- Process Modeling and Control - Creating a neural network model for a physical plant then using that model to determine the best control settings for the plant.
- Machine Diagnostics - Detect when a machine has failed so that the system can automatically shut down the machine when this occurs.
- Target Recognition - Military application which uses video and/or infrared image data to determine if an enemy target is present.
- Medical Diagnosis - Assisting doctors with their diagnosis by analyzing the reported symptoms and/or image data such as MRIs or X-rays.
- Financial Forecasting - Using the historical data of a security to predict the future movement of that security.
- Quality Control - Attaching a camera or sensor to the end of a production process to automatically inspect for defects.
- Intelligent Searching - An internet search engine that provides the most relevant content and banner ads based on the users' past behavior.

III. CONCLUSION

This paper provided a simplified approach to Neural Networks along with explanations on its different components. Neural Networks have gained so much ground that they are now termed as the sixth Generation of computing. As a matter of fact, Neural Networks have been applied in many fields such as science, finance, credit risk, economics and econometrics. Its ability to learn and being flexible render it a powerful tool though the black box problem reduces its usefulness. Nonetheless, the Predictive power of NN cannot be denied and this is making it still one of the best forecasting tools not Only among practitioners, let alone for central bankers in the world.

ACKNOWLEDGEMENT

We have immense pleasure in expressing our gratitude towards our paper guides Prof.S. S. Gawande, Prof D.R.Laddha for their time to time valuable guidance and inspiration given by them to us. They continuously supervised our work with utmost care. We also expressed our sincere thanks to Prof. P. L. Ramteke, H.O.D. of Information Technology, H.V.P.M's C.O.E.T.Amravati and also to all the teaching staff for their help and facilities provided for our paper work. We would like to thanks respected principal sir Dr. A. B. Marathe for providing necessary facility.

REFERENCES

[1] McCulloch, W.H., and Pitts, W.S., "A Logical Calculus of the Ideas Immanent in Neural Nets", Bulletin of Mathematical Biophysics, Vol. 5, 1943, pp. 115–133.

[2] The book **"Biology for Class XII:** NCERT".

[3] http://www.emilstefanov.net/Projects/NeuralNetworks.aspx.

[4] A.K. Jain, J. Mao, and K.M. Mohiuddin, "Artificial Neural Networks: A Tutorial", Computer, pp. 31-44, Mar, 1996.

[5] Intro to Neural Networks - Computer Vision Applications and Training Techniques. Doug Gray. www.soe.ucsc.edu/~taoswap/GroupMeeting/NN_Doug_2004_12_1.ppt.

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

[6] Introduction to Artificial Neural Networks. Nicolas Galoppo von Borries. www.cs.unc.edu/~nico/courses/ comp290-58/nn-presentation/ann-intro.ppt.

[7] The book "An introductory course on Neural Network "by Dr. Gunamani Jena and Dr. Rameswar Baliarsingh.

[8] N.B.Karayiannis, A.N.Venetsanopoulos, Artificial Neural Networks, Boston/Dordrecht/London:Kluwer Academic Publishers, 1993.

[9] Pioneering research into Brain Computer Interfaces Author: Mark Wessel, March 2006 Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science.]

[10] The book "Recent Advances In Artificial Neural Networks Design and Application" By Lakhmi Jain, Australia and Anna Maria Fanelli, Italy.

AUTHOR BIOGRAPHY

Ku. Swati G Anantwar , ME First Year(CSE). HVPM's COET AMRAVATI (MAHARASHTRA , INDIA)

Prof. R.R. Shelke; ME (CSE) LECTURER in HVPM's COET AMRAVATI (MAHARASHTRA , INDIA)