SciencePG
Science Publishing Group

An Overview of Neural Network

Mohaiminul Islam, Guorong Chen, Shangzhu Jin

School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing, China

Email address:
mohaiminul100@hotmail.com (M. Islam), szjin@cqust.edu.cn (Shangzhu Jin)

Abstract: Neural networks represent a brain metaphor for information processing. These models are biologically inspired rather than an exact replica of how the brain actually functions. Neural networks have been shown to be very promising systems in many forecasting applications and business classification applications due to their ability to learn from the data. This article aims to provide a brief overview of artificial neural network. The artificial neural network learns by updating the network architecture and connection weights so that the network can efficiently perform a task. It can learn either from available training patterns or automatically learn from examples or input-output relations. Neural network-based models continue to achieve impressive results on longstanding machine learning problems, but establishing their capacity to reason about abstract concepts has proven difficult. Building on previous efforts to solve this important feature of general-purpose learning systems, our latest paper sets out an approach for measuring abstract reasoning in learning machines, and reveals some important insights about the nature of generalization itself. Artificial neural networks can learn by example like the way humans do. An artificial neural net is configured for a specific application like pattern recognition through a learning process. Learning in biological systems consists of adjustments to the synaptic connections that exist between neurons. This is true of artificial neural networks as well. Artificial neural networks can be applied to an increasing number of real-world problems of considerable complexity. They are used for solving problems that are too complex for conventional technologies or those types of problems that do not have an algorithmic solution.

Keywords: Artificial Intelligence, Neural Network, Sigmoid Function, Neurons, Nodes

1. Introduction

Neural network is a type of artificial intelligence that attempts to imitate the way a human brain works. Rather than using a digital model, in which all computations manipulate zeros and ones, a neural network works by creating connections between processing elements, the computer equivalent of neurons. The organization and weights of the connections determine the output [1].

Neural networks are particularly effective for predicting events when the networks have a large database of prior examples to draw on. Strictly speaking, a neural network implies a non-digital computer, but neural networks can be simulated on digital computers.

Neural networks are a set of algorithms, modeled loosely after the human brain, that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated.

In its simplest form, a biological brain is a huge collection of neurons. Each neuron takes electrical and chemical signals as inputs through its many dendrites and transmits the output signals through its axon (in a more specialized context, there are exceptions to this behavior like with multipolar neurons). Axons make contact with other neurons at specialized junctions called synapses where they pass on their output signals to other neurons to repeat the same process over and over millions and millions of times [2].

Taking inspiration from the brain, an artificial neural network is a collection of connected units, also called neurons. The connection between the neurons can carry signals between them. Each connection carries a real number value which determines the weight/strength of the signal [3].
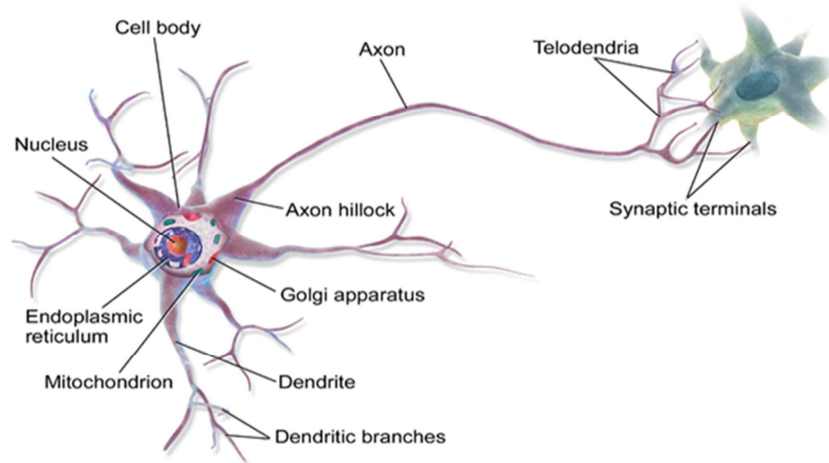
Figure 1. Neuron.

2. Types of Artificial Neural Network

Artificial neural networks are computational models which work similar to the functioning of a human nervous system. There are several kinds of artificial neural networks. These type of networks are implemented based on the mathematical operations and a set of parameters required to determine the output [4].

2.1. Feed Forward Neural Network Artificial Neuron

This neural network is one of the simplest form of ANN, where the data or the input travels in one direction. The data passes through the input nodes and exit on the output nodes. This neural network may or may not have the hidden layers. In simple words, it has a front propagated wave and no back propagation by using a classifying activation function usually.

Below is a Single layer feed forward network. Here, the sum of the products of inputs and weights are calculated and fed to the output. The output is considered if it is above a certain value i.e threshold (usually 0) and the neuron fires with an activated output (usually 1) and if it does not fire, the deactivated value is emitted (usually -1) [5].
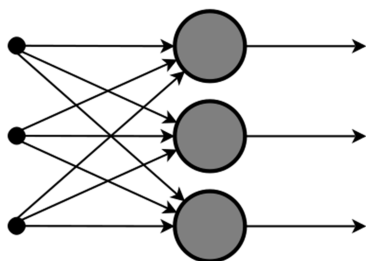

Figure 2. ANN.

Application of Feed forward neural networks are found in computer vision and speech recognition where classifying the target classes are complicated. These kinds of Neural Networks are responsive to noisy data and easy to maintain. This paper explains the usage of Feed Forward Neural Network. The X-Ray image fusion is a process of overlaying two or more images based on the edges.

2.2. Radial Basis Function Neural Network

Radial basic functions consider the distance of a point with respect to the center. RBF functions have two layers, first where the features are combined with the Radial Basis Function in the inner layer and then the output of these features are taken into consideration while computing the same output in the next time-step which is basically a memory.

Below is a diagram which represents the distance calculating from the center to a point in the plane similar to a radius of the circle. Here, the distance measure used in euclidean, other distance measures can also be used. The model depends on the maximum reach or the radius of the circle in classifying the points into different categories. If the point is in or around the radius, the likelihood of the new point begin classified into that class is high. There can be a transition while changing from one region to another and this can be controlled by the beta function.

This neural network has been applied in Power Restoration Systems. Power systems have increased in size and complexity. Both factors increase the risk of major power outages. After a blackout, power needs to be restored as quickly and reliably as possible. This paper how RBFnn has been implemented in this domain [6].

2.3. Kohonen Self Organizing Neural Network

The objective of a Kohonen map is to input vectors of arbitrary dimension to discrete map comprised of neurons. The map needs to me trained to create its own organization of the training data. It comprises of either one or two dimensions. When training the map the location of the neuron remains constant but the weights differ depending on the value. This self organization process has different parts, in the first phase every neuron value is initialized with a small weight and the input vector. In the second phase, the neuron closest to the point is the 'winning neuron' and the neurons connected to the winning neuron will also move towards the point like in the graphic below. The distance between the point and the neurons

is calculated by the euclidean distance, the neuron with the least distance wins. Through the iterations, all the points are clustered and each neuron represents each kind of cluster.

Kohonen Neural Network is used to recognize patterns in the data. Its application can be found in medical analysis to cluster data into different categories. Kohonen map was able to classify patients having glomerular or tubular with an high accuracy. Here is a detailed explanation of how it is categorized mathematically using the euclidean distance algorithm. Below is an image displaying a comparison between a healthy and a diseased glomerular [7].

2.4. Recurrent Neural Network (RNN) – Long Short Term Memory

The Recurrent Neural Network works on the principle of saving the output of a layer and feeding this back to the input to help in predicting the outcome of the layer.

Here, the first layer is formed similar to the feed forward neural network with the product of the sum of the weights and the features. The recurrent neural network process starts once this is computed, this means that from one time step to the next each neuron will remember some information it had in the previous time-step. This makes each neuron act like a memory cell in performing computations. In this process, we need to let the neural network to work on the front propagation and remember what information it needs for later use. Here, if the prediction is wrong we use the learning rate or error correction to make small changes so that it will gradually work towards making the right prediction during the back propagation.

2.5. Convolutional Neural Network

Convolutional neural networks are similar to feed forward neural networks, where the neurons have learn-able weights and biases. Its application has been in signal and image processing which takes over OpenCV in field of computer vision.

ConvNet are applied in techniques like signal processing and image classification techniques. Computer vision techniques are dominated by convolutional neural networks because of their accuracy in image classification. The technique of image analysis and recognition, where the agriculture and weather features are extracted from the open source satellites like LSAT to predict the future growth and yield of a particular land are being implemented [8].

2.6. Modular Neural Network

Modular Neural Networks have a collection of different networks working independently and contributing towards the output. Each neural network has a set of inputs which are unique compared to other networks constructing and performing sub-tasks. These networks do not interact or signal each other in accomplishing the tasks. The advantage of a modular neural network is that it breakdowns a large computational process into smaller components decreasing the complexity. This breakdown will help in decreasing the number of connections and negates the interaction of these

networks with each other, which in turn will increase the computation speed. However, the processing time will depend on the number of neurons and their involvement in computing the results.

3. Artificial Neural Network (ANN) Algorithm Work

A typical neural network has anything from a few dozen to hundreds, thousands, or even millions of artificial neurons called units arranged in a series of layers, each of which connects to the layers on either side. Some of them, known as input units, are designed to receive various forms of information from the outside world that the network will attempt to learn about, recognize, or otherwise process. Other units sit on the opposite side of the network and signal how it responds to the information it's learned; those are known as output units. In between the input units and output units are one or more layers of hidden units, which, together, form the majority of the artificial brain. Most neural networks are fully connected, which means each hidden unit and each output unit is connected to every unit in the layers either side. The connections between one unit and another are represented by a number called a weight, which can be either positive (if one unit excites another) or negative (if one unit suppresses or inhibits another). The higher the weight, the more influence one unit has on another. (This corresponds to the way actual brain cells trigger one another across tiny gaps called synapses [9].

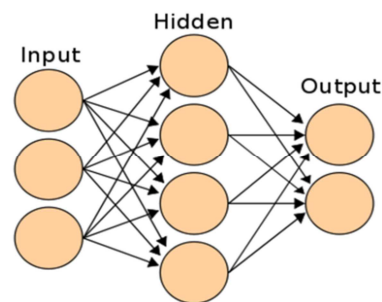### 3.1. Formulation of Neural Network: A Simple Neural Network can Be Represented as Shown in the Figure Below

Figure 3. Input - Hidden layer – Output.

The linkages between nodes are the most crucial finding in an ANN. We will get back to "how to find the weight of each linkage" after discussing the broad framework. The only known values in the above diagram are the inputs. Lets call the inputs as I1, I2 and I3, Hidden states as H1, H2, H3 and H4, Outputs as O1 and O2 [7]. The weights of the linkages can be denoted with following notation:

W(I1H1) is the weight of linkage between I1 and H1 nodes.

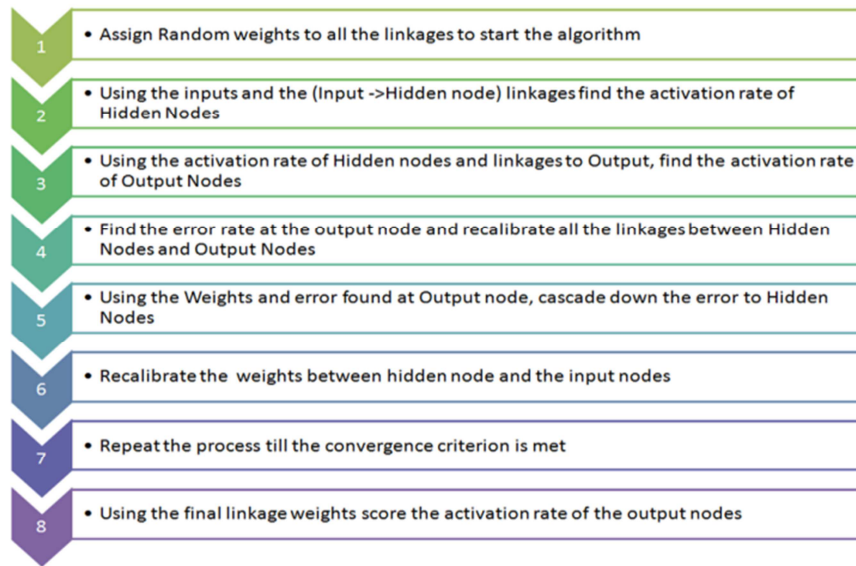Following is the framework in which artificial neural networks (ANN) work:

Figure 4. *Work process.*

3.2. Few Statistical Details about the Framework

Every linkage calculation in an Artificial Neural Network (ANN) is similar. In general, we assume a sigmoid relationship between the input variables and the activation rate of hidden nodes or between the hidden nodes and the activation rate of output nodes. Let's prepare the equation to find activation rate of H1.

Logit (H1) = W(I1H1) * I1 + W(I2H1) * I2 + W(I3H1) * I3 + Constant = f

=> P(H1) = 1/(1+e^(-f))

Following is how the sigmoid relationship looks like:
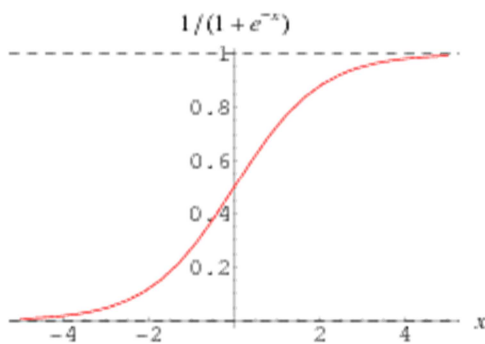

Figure 5. *Sigmoid function".*

3.3. The Weights re-Calibrated

Re-calibration of weights is an easy, but a lengthy process. The only nodes where we know the error rate are the output nodes. Re-calibration of weights on the linkage between hidden node and output node is a function of this error rate on output nodes. It can be statistically proved that:

Error @ H1 = W(H1O1)*Error@O1 + W(H1O2)*Error@O2

Using these errors we can re-calibrate the weights of linkage between hidden nodes and the input nodes in a similar fashion. Imagine, that this calculation is done multiple times for each of the observation in the training set [10].

4. Backpropagation

Back propagation is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network. Back propagation is shorthand for "the backward propagation of errors," since an error is computed at the output and distributed backwards throughout the network's layers. It is commonly used to train deep neural networks.

Back propagation is a generalization of the delta rule to muti-layered feed forward networks, made possible by using the chain rule to alliterative compute gradients for each layer. It is closely related to the Gauss–Newton algorithm and is part of continuing research in neural back propagation.

Back propagation is a special case of a more general technique called automatic differentiation. In the context of learning, back propagation is commonly used by the gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function [11].

4.1. ANN and DNN Concepts Relevant to Backpropagation

Here are several neural network concepts that are important to know before learning about backpropagation:
Inputs
Source data fed into the neural network, with the goal of making a decision or prediction about the data. The data is broken down into binary signals, to allow it to be processed by single neurons—for example an image is input as individual pixels.

4.2. Training Set

A set of outputs for which the correct outputs are known, it

can be used to train the neural networks.

Outputs

The output of the neural network can be a real value between 0 and 1, a boolean, or a discrete value (for example, a category ID).

4.3. Activation Function

Each neuron accepts part of the input and passes it through the activation function. Commonly used functions are the sigmoid function, tanh and ReLu. Modern activation functions normalize the output to a given range, to ensure the model has stable convergence [12].

4.4. Weight Space

Each neuron is given a numeric weight. The weights, applied to the activation function, determine each neuron's output. In training of a deep learning model, the objective is to discover the weights that can generate the most accurate output.

4.5. Initialization

Setting the weights at the beginning, before the model trained, a typical strategy in neural networks is to initialize the weights randomly, and then start optimizing from there. Xavier optimization is another approach which makes sure weights are "just right" to ensure enough signal passes through all layers of the network13.

4.6. Forward Pass

The forward pass tries out the model by taking the inputs, passing them through the network and allowing each neuron to react to a fraction of the input, and eventually generating an output.

4.7. Gradient Descent

A mathematical technique that modifies the parameters of a function to descend from a high value of a function to a low value, by looking at the derivatives of the function with respect to each of its parameters, and seeing which step, via which parameter, is the next best step to minimize the function. Applying gradient descent to the error function helps find weights that achieve lower and lower error values, making the model gradually more accurate [14].

5. 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 respond computational times which are due to their parallel architecture.

Neural networks also contribute to other areas of research such as neurology and psychology. They are regularly used to model parts of living organisms and to investigate the internal mechanisms of the brain.

Perhaps the most exciting aspect of neural networks is the possibility that some day 'consious' networks might be produced. There is a number of scientists arguing that conciousness is a 'mechanical' property and that 'consious' neural networks are a realistic possibility [15].

In future, If anyone want to develop this research, he can choose a topic from here. This paper is an overview of artificial neural network. It's better to choose a small topic and do research on it.

References

[1] Neural Networks at Pacific Northwest National Laboratory http://www.emsl.pnl.gov:2080/docs/cie/neural/neural.homepage.html

[2] Klimasauskas, CC. (1989). The 1989 Neuro Computing Bibliography. Hammerstrom, D. (1986). A Connectionist/Neural Network Bibliography.

[3] N. Murata, S. Yoshizawa, and S. Amari, ―Learning curves, model selection and complexity of neural networks,‖ in Advances in Neural Information Processing Systems 5, S. Jose Hanson, J. D. Cowan, and C. Lee Giles, ed. San Mateo, CA: Morgan Kaufmann, 1993, pp. 607-614.

[4] Bradshaw, J. A., Carden, K. J., Riordan, D., 1991. Ecological ―Applications Using a Novel Expert System Shell‖. Comp. Appl. Biosci. 7, 79–83.

[5] Lippmann, R. P., 1987. An introduction to computing with neural nets. IEEE Accost. Speech Signal Process. Mag., April: 4-22.

[6] Murphy, K. P. *Machine Learning: A Probabilistic Perspective.* Cambridge, Massachusetts: The MIT Press, 2012.

[7] Hubel, H. D. and Wiesel, T. N. " Receptive Fields of Single neurones in the Cat's Striate Cortex." Journal of Physiology. Vol 148, pp. 574-591, 1959.

[8] Jeannette Lawrence, *Data Preparation for a Neural Network*, Neural NetworkSpecial Report: A Miller Freeman Publication, 1992.

[9] Gene Bylinsky, Computers That Learn By Doing, Fortune, September 6, 1993.

[10] Dennis Collins, Brain Maker: Strange, Captivating, Easy to Use, California Computer News, July, 1990.

[11] 1993 Readers' Choice Awards Winner: Brain Maker Professional, Software: Artificial Intelligence, Technical Analysis of Stocks & Commodities, Bonus Issue 1994.

[12] Carlton F. Vogt Jr., Brain Maker: This Will Teach Your Computer to Think, Design News, December 3, 1990.

[13] John Vester, Artificial Intelligence and Real Life, Computor Edge, October 22, 1993.

[14] George W. Dombi and Jeannette Lawrence, Analysis of protein transmembrane helicalregions by a neural network, Protein Science (1994), 3:557-566.

[15] Henrik Lundstedt, Neural Networks and Predictions of Solar-terrestrial Effects, Lund Observatory, Lund, Sweden, 1990.