www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 5 Issue 11 Nov. 2016, Page No. 18965-18971

Utility Study of Neural Network and Back Propagation Algorithm in The Field Of Learning And Computing Data In Mines Area

Prof. Amar Nath Singh[1] Er. Amit Kumar Patra[2], Er. Biswabijayee Dash[2], Er. MD Aqueeb Jawed[2] Er. Krishnananda Nayak[2]
Gandhi Engineering College,
Bhubaneswar, Odisha
Email: amarnathsingh.2k@gmail.com
Gandhi Engineering College, Bhubaneswar, Odisha

ABSTRACT

Neural Networks (NN) are important data mining (Extraction of Knowledge) tool used for classification and clustering (Grouping of nodes together). It is a branch of Artificial Intelligence which plays an important role in almost all field of science. It is an attempt to build machine that will mimic brain activities and be able to learn. NN usually learns by examples. If NN is supplied with enough examples, it should be able to perform classification and even discover new trends or patterns in data. If we consider the architecture point of view then, NN is composed of three layers, such as input, output and hidden layer. Each layer can have number of nodes and nodes from input layer are connected to the nodes from hidden layer. Nodes from hidden layer are connected to the nodes from output layer. Those connections represent weights between nodes.
The Back Propagation (BP) Algorithm is one of most popular NN algorithms, which is applied in every sector of real time application. BP algorithm is quite simple, eases to handle and works on the principle of, output of NN, which is evaluated against desired output. If results are not satisfactory, connection (weights) between layers are modified and process is repeated again and again until error is small enough. In underground mine area, we apply this algorithm to determine the valuable data like, signal analysis, rock characterization, etc.
Keywords: Neural Networks, Artificial Neural Networks, Back Propagation algorithm, Underground Mine, Coal Mine.

1 Introduction

Classification is grouping of the objects or things that are similar. Examples of classifications are found everywhere, supermarkets put similar things together, and there are shelves for meat, dairy products, cleaning products.
If we take the example of book shelves then the arrangement of book in the shelf make the searching of a book easier. Similarly in most of the cases the arrangement is highly required. There are many classification techniques used in data mining with NN being one of them. NN model is Explained in this paper but distinction has to be made between NN in humans and animals and NN that are being used in industries[1]. Previously the Neural Network are often called Artificial NN but for simplicity they will be referred to as NN.
According to the classification point of view, there are two types of NN based on learning technique, they are as follows, such as

1. Supervised NN: It is a NN architecture, where output values are known before hand (back propagation algorithm) and,
2. Unsupervised NN: It is the another form of NN, where output values are not known (clustering). They have to first computed and then admired.

In the NN architecture, number of nodes to choose how to set the weights between the nodes, training the network and evaluating the results are covered. Activation function gets mentioned together with learning
Rate, momentum and pruning. Back propagation algorithm, probably the most popular NN algorithm is also discussed here [1, 2].

2 Neural Networks

DOI: 10.18535/ijecs/v5i11.42

Neural networks have seen an explosion of interest over the last few years and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology and physics.' Statsoft.com [2010].
It was first started in 1943 when McCullock and Pitts proved that neuron can have two states and that those states could be dependent on some threshold value. They presented first artificial neuron model according to Rojas [2005]. Then after many other model are came into the picture and many new and more sophisticated models have been presented since.

According to the Muller et al. [1995] there are two main reason for NN investigation,

1. First is to try to get an understanding on how human brain function and,
2. Second is desire to build machines that are capable for solving complex problems that sequentially operating computers were unable to solve.

So, if problem structure is well analyzed, traditional computers could still outperform NN but in cases where problem has not been analyzed in details, NN could be used to learn from large set of examples. NN can handle errors better than traditional computers programs (imagine scenario where one faulty statement in program could halt everything while NN can handle errors in much better manner).

2.1 Neuron

The neuron is nothing but a human living cell. According to the scientist Vaga [1994] Human brain has over 100 billion interconnected neurons. Most sophisticated application have only tiny fraction of that. It can only be imagined how powerful NN with this number of interconnected neurons would be.
The Neurons use this interconnected network to pass information's with each other using electric and chemical signals. Although it may seem that neurons are fully connected, two neurons actually do not touch each other. They are separated by tiny gap call Synapse. Each neuron process information and then it can "connect" to as many as 50 000 other neurons to exchange information. If connection between two neuron is strong enough (will be explained later) information will be passed from one neuron to another.
On their own, each neuron is not very bright but put 100 billion of them together and let them talk to each other, then this system becomes very powerful[5].

In terms of structural classification point of view, A typical neuron would have 4 components seen on Dendrites, Soma, Axon and Synapse. Dendrites gather inputs from other neurons and when a certain threshold is reached they generate a non-linear response (signal to other neurons via the Axon).

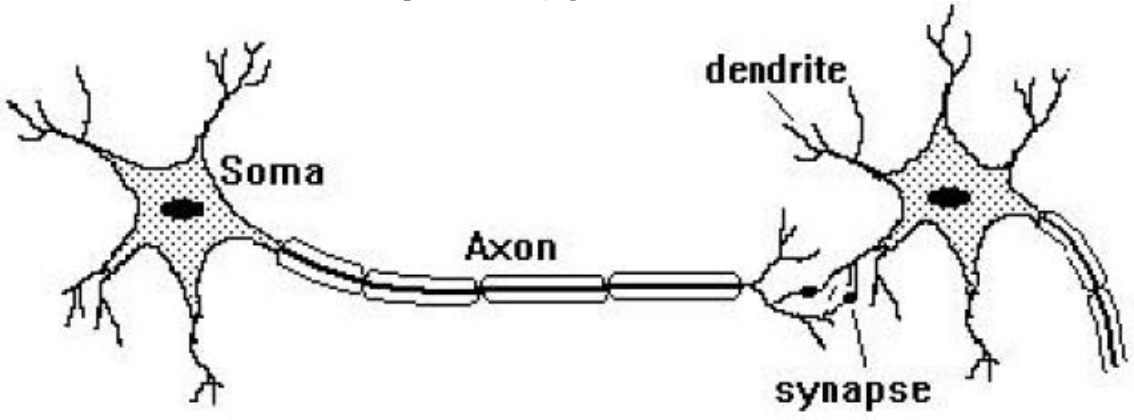


Figure 1: ComputingAssociate.com [2010]

2.2 Architecture

The figure showed below, represent architecture of an simple NN. Usually the NN, made up from an input, output and one or more hidden layers. Here in the structure, each node from input layer is connected to a node from hidden layer and every node from hidden layer is connected to a node in output layer. The main purpose of the Input layer is to represents an the raw information that is fed into the network[5]. This part of network is never changing its values. Hence it is also sometimes refer as the fixed structure node. Every single input to the network is duplicated and send down to the nodes in hidden layer. Hidden Layer accepts data from the input layer. It uses input values and modifies them using some weight value. The new value so obtained is than send to the output layer but it will also be modified

by some weight from connection between hidden and output layer. And in this way the nodes are going to be operated in the structure.

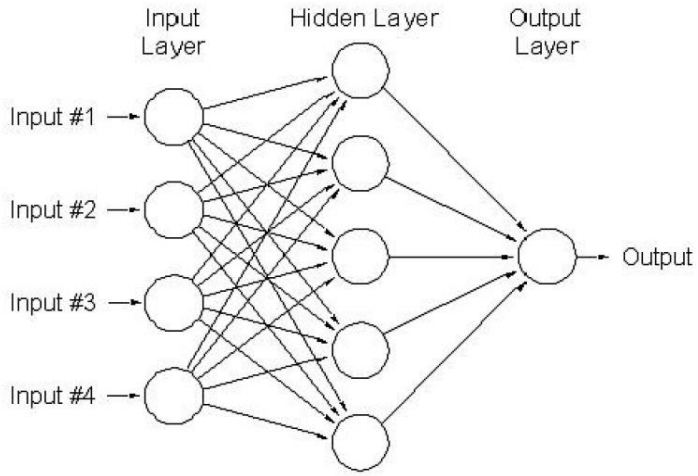


Figure 2: Simple Neural Network

2.3 Number of Nodes and Layers

Choosing number of nodes for each layer will depend on problem NN is trying to solve, types of data network is dealing with, quality of data and some other parameters. Number of input and output nodes depends on training set in hand. Larose [2005] argued that choosing number of nodes in hidden layer could be challenging task. If there are too many nodes in hidden layer, number of possible computations that algorithm has to deal with increases[5]. Picking just few nodes in hidden layer can prevent the algorithm of its learning ability.
Right balance needs to be picked. It is very important to monitor the progress of NN during its training, if results are not improving, some modification to the model might be needed.

2.4 Setting Weights

As we know that, in the given NN arrangement, the weight factor is also associated. The way to control NN is by setting and adjusting weights between nodes. Initial weights are usually set at some random numbers and than they are adjusted during NN training.
According to Fogel [2002] focus should not be at changing one weight at time, changing all the weights should be attempted simultaneously. Some NN are dealing with thousands, even millions of nodes so changing one or two at time would not help in adjusting NN to get desired results in timely manner.
Logic behind weight updates is quite simple.

2.5 Running and training NN

Running the network consist of a forward pass and a backward pass. In the forward pass outputs are calculated and compared with desired outputs. Error from desired and actual output are calculated. In the backward pass this error is used to alter the weights in the network in order to reduce the size of the error. Forward and backward pass are repeated until the error is low enough (users usually set the value of accepted error)[4, 5].

2.6 Activation Function

Activation function is the simple linear function. They can be linear, threshold or sigmoid function. Sigmoid activation function is usually used for hidden layer because it combines nearly linear behavior, curvilinear behavior and nearly constant behavior depending on the input value Larose [2005]. To explain activation function figure 3 will be used which is as showed bellow.

DOI: 10.18535/ijecs/v5i11.42

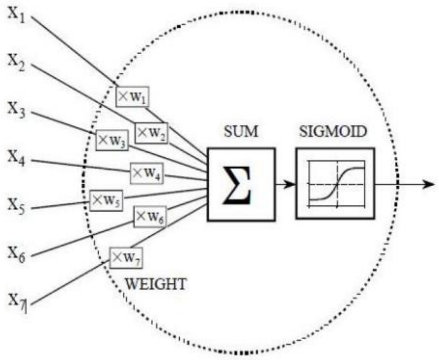


Figure 3: Dspguide.com [2010]

SUM is collection of the output nodes from hidden layer that have been multiplied by connection weights, added to get single number and put through sigmoid function (activation function). Input to sigmoid is any value between negative infinity and positive infinity number while the output can only be a number between 0 and 1.

2.7 Pruning

Pruning could be useful technique when dealing with large NN. According to Michie et al. [1994] pruning useless nodes or weights have numerous advantages on NN. Smaller networks and faster training times on serial computers have some advantages. Smaller networks are easier to interpret but pruning should be considered carefully as it can impact network performance.

3. Back Propagation (BP) Algorithm

One of the most popular NN algorithms is back propagation algorithm. Rojas [2005] claimed that BP algorithm could be broken down to four main steps. After choosing the weights of the network randomly, the back propagation algorithm is used to compute the necessary corrections. The algorithm can be decomposed in the following four steps:

i) Feed-forward computation
ii) Back propagation to the output layer
iii) Back propagation to the hidden layer
iv) Weight updates

The algorithm is stopped when the value of the error function has become sufficiently small. This is very rough and basic formula for BP algorithm. There are some variations proposed by other scientist but Rojas definition seem to be quite accurate and easy to follow.

3.1 Worked example

To demonstrate the concept of Back propagation technique, let us consider a example in which we have a NN , which has two nodes (N0,0 and N0,1) in input layer, two nodes in hidden layer (N1,0 and N1,1) and one node in output layer (N2,0).

Now according to BP (Back Propagation) to compute the factor, let us consider the weight associated with each node. We have already depicted the node here, so let us assume that, the Input layer nodes are connected to hidden layer nodes with weights (W0,1-W0,4). Hidden layer nodes are connected with output layer node with weights (W1,0 and W1,1). The values that were given to weights are taken randomly and will be changed during BP iterations.
Refer to figure-4 showed bellow.

3.1.1 Feed-forward computation

It is a first classification of BP where Feed forward computation or forward pass is two step process. First part is getting the values of the hidden layer nodes and second part is using those values from hidden layer to compute value or values of output layer. Input values of nodes N0,0 and N0,1 are pushed up to the network towards nodes in hidden

layer (N1,0 and N1,1). They are multiplied with weights of connecting nodes and values of hidden layer nodes are calculated.

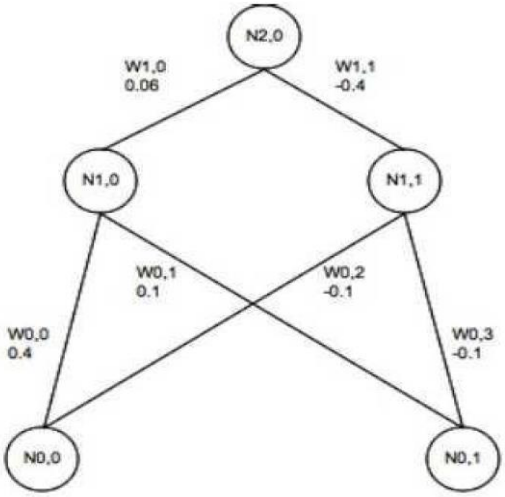
Here the Sigmoid function is used for calculations f(x) = 1.0=(1.0 + exp(-x)).

N1,0 = f(x1) = f(w0,0* n0,0 + w0,1*n0,1) = f(0.4 + 0.1) = f(0.5) = 0.622459
N1,1 = f(x2) = f(w0,2 * n0,0 + w0,3* n0,1) = f(-0.1 – 0.1) = f(-0:2) = 0.450166

When hidden layer values are calculated, network propagates forward, it propagates values from hidden layer up to a output layer node (N2,0). This is second step of feed forward computation

N2,0 = f(x3) = f(w1,0* n1,0 + w1,1 * n1,1) = f(0.06* 0:622459 + (-0:4)* 0:450166)
= f(-0.1427188) = 0.464381

Having calculated N2,0, forward pass is completed.



Pattern data for AND

n0,0	*n0,1*	*Output n2,0*
1	1	1
1	0	0
0	1	0
0	0	0

β = Learning rate = 0.45
α = Momentum term = 0.9

$f(x) = 1.0 / (1.0 + \exp(-x))$

Figure 4: Example

3.1.2 Back propagation to the output layer

It is the Next step to calculate error of N2,0 node. From the table in figure 4, output should be 1. Predicted value (N2,0) in our example is 0.464381. Error calculation is done the following way:

N2,0Error = n2,0*(1-n2,0)*(N2,0Desired - N2,0) = 0.464381(1-0.464381)*(1- 0.464381) = 0.133225

Once error is known, it will be used for backward propagation and weights adjustment. It is two step process. Error is propagated from output layer to the hidden layer first. This is where learning rate and momentum are brought to

equation. So weights W1,0 and W1,1 will be updated first. Before weights can be updated, rate of change needs to be found. This is done by multiplication of the learning rate, error value and node N1,0 value.

ΔW1,0 =β* N2,0Error * n1,0 = 0.45 * 0.133225* 0.622459 = 0.037317

Now new weight for W1,0 can be calculated.

W1,0New = w1,0Old + ΔW1,0 + (α * ∇t - 1)) = 0.06 + 0.037317 + 0.9 * 0 = 0.097137
ΔW1,1 =β* N2,0Error * n1,1 = 0.45 * 0.133225* 0.450166 = 0.026988
W1,1New = w1,1Old + ΔW1,1 + (α * ∇t -1)) = -0.4 + 0.026988 = -0:373012

3.1.3 Back propagation to the hidden layer

This is bit more complicated than propagating error from output to hidden layer. In previous case, output from node N2,0 was known before hand. Output of nodes N1,0 and N1,1 was unknown. This will be calculated multiplying new weight W1,0 value with error for the node N2,0 value. Same way error
for N1,1 node will be found.

N1, 0Error = N2,0 Error *W1, 0New = 0.133225* 0.097317 = 0.012965
N1, 1Error = N2, 0Error *W1,1New = 0.133225 * (-0.373012) = -0.049706
Once error for hidden layer nodes is known, weights between input and hidden layer can be updated.

3.1.4 Weight updates

Important thing is not to update any weights until all errors have been calculated. It is easy to forget this
and if new weights were used while calculating errors, results would not be valid. Here is quick second pass using new weights to see if error has decreased.

N1,0 = f(x1) = f(w0, 0 * n0,0 + w0,1 * n0,1) = f(0.406 + 0.1) = f(0.506) = 0.623868314
N1,1 = f(x2) = f(w0,2 * n0,0 + w0,3 * n0,1) = f(-0.122 – 0.122) = f(-0.244) = 0.43930085
N2,0 = f(x3) = f(w1, 0 * n1,0 + w1,1 * n1,1) = f(0.097 * 0.623868314 + (-0:373) * 0.43930085) = f(-0.103343991) = 0.474186972

Having calculated N2,0, forward pass is completed.

4. Advantages and Disadvantages

We know that the NN is a case of Artificial Intelligence, so in this case we need to calculate the factor associated with the node. In 2005, a scientist Negnevitsky argued that turning point in quest for intelligent machines was when Kasparov, world chess champion was defeated by computer in New York in May 1997.
Artificial intelligence and NN have been used more and more in recent decades. Because it is a more advance research area where we can compute the intelligence.
If we look to advantage and disadvantages, then NN are used in cases where rules or criteria for searching an answer is not clear (that is why NN are often called black box, they can solve the problem but at times it is hard to explain how problem was solved)[6, 7].
They found its way into broad spectrum of industries, from medicine to marketing and military just to
name few. Financial sector has been known for using NN in classifying credit rating and market forecasts.
Marketing is another field where NN has been used for customer classification (groups that will buy some
product, identifying new markets for certain products, relationships between customer and company). Many companies use direct marketing (sending its over by mails) to attract customers. If NN could be employed to up the percentage of the response to direct marketing, it could save companies lot's of their revenue. At the end of the day, it's all about the money.
They also require very large sample sets to train model efficiently. It is hard to explain results and what is going on inside NN.

5. Conclusion

Neural Network is an application of Artificial intelligence which is interconnected network that resembles human brain. The most important characteristic of NN is its ability to learn.
We know that, the NN is classified into Supervised NN: It is a NN architecture, where output values are known before hand (back propagation algorithm) and Unsupervised NN: It is the another form of NN, where output values are not known (clustering). They have to first computed and then admired. Whatever be the cases may be, the importance of AI along with the NN is presented with training set (form of supervised learning) where input and output values are known, NN model could be created to help with classifying new data. It is a vast research area which imparts a great role in modern science and Engineering. It is also plays a vital role in modern medical science and in the field of underground and surface mining. Mining is the most important application for now a day, so we can also apply the BP algorithms for the field of coal mining. In such field the human interaction with the machine is highly essential which can be made possible due to advance in Artificial Intelligence as well as NN? It also utilizes in pattern recognition. NN is getting more and more attention in last two decades. BP algorithm is most popular algorithm used in NN. It is one of the main reasons why NN are becoming so popular.

7. Acknowledgement

Prof. Amar Nath Singh, is a Reader in the department of Computer Science and Engineering, GEC, Bhubaneswar, Odisha. He received his master degree in the year of 2007 form BPUT, Rourkela, Odisha, India. He is perusing his PhD in the field of Mines area. His research are includes Underground Mines, ArtificialIntelligence, Wireless Sensor Network and Expert system, Fuzzy Logic Network, HCL, Algorithm, Web logic, etc.

6. References

1. ComputingAssociate.com. Introduction to Neural Networks. Retrived from internet on 26 Oct 2010 http://euclid.ii.metu.edu.tr/ ion526/demo/chapter1/section1.1/_gure11e:html; 2010:
2. Dspguide.com. Introduction to Neural Networks. Retrived from internet on 30 Oct 2010 http://www.dspguide.com/CH26.PDF, 2010.
3. Faqs.org. Why use activation functions. Retrived from internet on 30 Oct 2010 http://www.faqs.org/faqs/ai-faq/neural-nets/part2/section-10.html, 2010.
4. D.B. Fogel. Blondie24: playing at the edge of AI. The Morgan Kaufmann Series in Evolutionary Computation. Morgan Kaufmann Publishers, 2002. ISBN 9781558607835. URL http://books.google.ie/books?id=5vpuw8L0CAC: \
5. D.T. Larose. Discovering knowledge in data: an introduction to data mining. Wiley-Interscience, 2005. ISBN 9780471666578. URL http://books.google.ie/books?id=JbPMdPWQIOwC.
6. D. Michie, D.J. Spiegelhalter, and C.C. Taylor. Machine learning, neural and statistical classi_cation. Arti_cial intelligence. Ellis Horwood, 1994. ISBN 9780131063600. URL http://books.google.ie/books?id=QPqCQgAACAAJ.
7. B. Muller, J. Reinhardt, and M.T. Strickland. Neural networks: an introduction. Number v. 1 in Physics of neural networks. Springer, 1995. ISBN 9783540602071. URL http://books.google.ie/books?id=EFUzMYjOXk8C.