Univariate Analysis for Rolling Mill Using Artificial Neural Network

Ajay Batra
Mechanical Engineering Department,
YITM, Rajnandgaon, CSVTU,
Chhattisgarh, India.

cArvind Rai
Mechanical Engineering Department
BIT, Durg, CSVTU,
Chhattisgarh, India
batra2106@yahoo.com

ABSTRACT
The behavior of any company is complicated. This work aims at evaluation of financial performance for business specific company. Numerous financial indicators can be tabled; but bear sparing relation to output performance.

Several efforts were made by scientists to bind input output parameters. Neural network is similar technique to grip input output binder. It handles any non-linearity with an ease.

This work employs Artificial Neural Network based model to estimate profit using four independent parameters for Rolling Mill. The output so obtained has less error, compared to Regression Analysis. Established relation permits to understand detailed intricate behavior of the sector and thereby analyze the criticality of parameters under consideration. It may be useful to Companies, Board Members, Stock Holders, and Entrepreneurs.

Keywords
Correlation coefficient, Neural networks, Back Propagation, Univariate Analysis.

1. INTRODUCTION
All industries operate and achieve their goal by using a system in which a set of inputs are fed to achieve desired outputs over a period of time. There exists a relation between these input and output variables. Management can be aided with powerful decision making if a model establishing relation as well as criticality of these variables can be developed. Hence by addressing this very need, a model is devised for the given problem? At present, rolling mill is considered for study.

1.1 Problem Definition & Objective Function
"To design a model which can establish a relationship between the inputs and output variables of an industrial sector (rolling mill) in order to predict the output & analyze the influence of different input parameters, thereby establishing their critically and the length of influence of that particular parameter."

This project consists of an evaluation cum forecasting model which can establish a relationship between the input and output of an industrial sector. The model is based on the interrelationship of the inputs and outputs which would be studied as well as corrected to a considerable extent so as aid in predicting and analyzing performance of the industry.

The model will enlighten the existing companies as well as the new entrants in the specified sector. An artificial neural network is a system based on the operation of biological neural networks which is an adaptive, most often nonlinear system that learns to perform a function (an input/output map) from data. Correlation and Vogel's approximation method are used to justify input parameter selection for artificial neural network. These parameters are changed during operation, normally called training phase of the artificial neural network. After the training phase the Artificial Neural Network parameters are fixed and the system is deployed to solve the problem at hand.

To perform Univariate analysis to find the most influential input parameter and length of influence of that particular which affects the output.

1.2 Available Methods and Selection of Best Alternative
1. Regression Analysis
2. Artificial Neural Network (ANN) along with correlation

Reasons for Selecting ANN as the Best Alternative

1. The forecasted results obtained as error in the subsequent regression analysis were significantly high.
2. A neural network can perform tasks that a linear program cannot.
3. When an element of the neural network fails, it can continue without any problems by their parallel nature.
4. A neural network learns and does not need to be reprogrammed.
5. Back Propagation method has been found to be a strong forecasting tool and is in an advantageous position in comparison to Regression analysis [1].

1.3 Artificial Neural network (ANN)
Artificial neural network is an information processing model based upon the function of neurons or nerve cells, found in human brain and nervous system [2]. Thus, first of all, we need to consider the essential properties of biological neural networks from the viewpoint of information processing. This will allow us to design abstract models of artificial neural networks, which can then be simulated and analyzed.

Although the models which have been proposed to explain the structure of the brain and the nervous systems of some animals are different in many respects, there is a general consensus that the essence of the operation of neural ensembles is "control through communication". Animal nervous systems are composed of thousands or millions of interconnected cells. Each one of them is a very complex arrangement which deals with incoming signals in many different ways. However, neurons are rather slow when compared to electronic logic gates. These can achieve switching times of a few nanoseconds whereas neurons need several milliseconds to react to a stimulus. Nevertheless the brain is capable of solving problems which no digital computer can yet efficiently deal with. But then a significant achievement can be achieved with the help of Artificial Neural Networks.

Artificial Neural Networks have aroused so much interest in recent years, not only because they exhibit interesting properties, but also because they try to mirror the kind of information processing capabilities of nervous systems. Artificial Neural Network (ANN) is a system based on the operation of biological neural networks, in other words, it is an emulation of biological neural system. Although computing these days is truly advanced, there are certain tasks that a program made for a common microprocessor is unable to perform; even so a software implementation of a neural network can be made. An Artificial Neural Network is an adaptive, most often nonlinear system that learns to perform a function (an input/output map) from data. Adaptive means that the system parameters are changed during operation, normally called the training phase. After the training phase the Artificial Neural Network parameters are fixed and the system is deployed to solve the problem at hand (the testing phase). Important characteristic that ANN share with biological neural system is fault tolerance. The NNs can be designed to be insensitive to small damage to the network and the network can be retained in cases of significant damage [3].

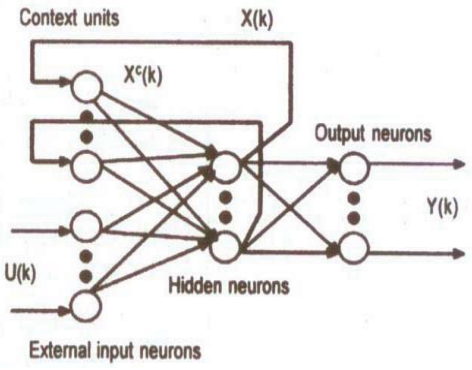

Fig.1 Elman Network

As shown in figure 1, the standard Elman network consists of input/output and context units. The input/output units behave similarly as in feed-forward networks i.e. the input units serve as buffers to distribute the signals without processing them, and, output units linearly sum the inputs from the preceding layer and have a linear activation function. The hidden units can have linear or non linear activation function. Contexts units are used to memorize previous activations of the hidden units and acts as one step time delays [4].

Analysis of Elman network is provided by Pham [5]. At a given time k, the input to the network consist of the current input u(k) and the previous activations of the hidden units x(k-1). These inputs are propagated forward to produce the outputs. Standard back-propagation algorithm can then be used to train the network. Activation of the hidden units at time k are sent back through the recurrent links to the context units and saved there for the next training step at time (k+1). External inputs to the network are represented by U(k-1) and the network outputs by Y(k). Activation of the hidden units are represented by $X^c(k)$. Elman network is able to model an n^{th} order dynamic system (where n is the number of units in the hidden layer) if it can be trained to do so. For that 2n input units would be needed if tapped delay line method is used. For an Elman network, the number of units is one, or n+1 if the context units are regarded as inputs units, hence it is significantly smaller in structure then a feed-forward when n is large.

In real life problems we have a number of input parameters. The relationship between input and output parameters is usually complex. Therefore we need some mathematical tool in order to establish the relationship so that the output obtained would mirror the original output. Therefore correlation analysis has been used into find out the interdependencies of variables.

1.4 ANN along with "Correlation"

Correlation is used to show the dependency of the input variables on each other. In probability theory and statistics correlation (often measured as correlation coefficient) indicates the strength and direction of a linear relationship between two random variables. Correlation is a statistical technique that can show whether and how strongly pairs of variables are related. The main result of a correlation is called the correlation coefficient.

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

It ranges from -1.0 to +1.0. The closer r is to +1 or -1, the more closely the two variables are related. If r is close to 0, it means there is no relationship between the variables. If r is positive, it means that as one variable gets larger, the other gets larger. If r is negative, it means that as one variable gets larger, the other gets smaller (often called an "inverse" correlation). The correlation matrix of n random variables X_1, X_2, ……., X_n is the n x n matrix whose i, j entry is corr (X_i, X_j). If the measures of correlation used are product-moment coefficients, the correlation matrix is the same as the covariance matrix of the standardized random variables X_i/SD (X_i) for i= 1,…, n. consequently it is necessarily a positive semi definite matrix. The correlation matrix is symmetric because the correlation X_i and X_j is the same as the correlation between X_j and X_i.

There are many factors responsible for running an industry. But here a very genuine question arises that does each factor which plays a significant role in itself dependent on other factor or totally independent of each other. To prove that there exists some relationship between each factor, we need a mathematical tool. Correlation is that statistical tool which helps us to find a definite relationship between a numbers of factors. Through correlation we come to know that how much a factor in the industry is dependent or what kind of relationship it possesses. Correlation coefficient is denoted by r which ranges from -1 to +1. If the value is near +1 or -1 then the factors are closely related. And if it is 0 then there is no relationship between variables. Now, selection of important variables needs to be done.

1.5 Learning Algorithms for Neural Networks

A learning algorithm is an adaptive method by which a network of computing units self-organizes to implement the desired behavior. This is done in some learning algorithms by presenting some examples of the desired input output mapping to the network. Haykin [6] outlines how neural networks mimic a very simplified version of the brain in two aspects. Neurons are the processing elements of the network. The layout of the neurons in space and the interconnections between them determine the structure of the network. Common structures can be modified to suit a particular application, or alternatively a completely new structure can be designed. A learning algorithm decides how the weights are adjusted to achieve the desired behavior of the network. A correction step is executed iteratively until the network learns to produce the desired response. The learning algorithm is a closed loop of presentation of examples and of corrections to the network.

1.6 Classes of Learning Algorithms

Learning algorithm can be divided into supervised and unsupervised methods [7].

Supervised learning denotes a method in which some input vectors are collected and presented to the network. The output computed by the network is observed and the deviation from the expected answer is measured. The weights are corrected according to the magnitude of the error in the way define by the learning algorithm. This kind of learning is also called learning with teacher, since a control process knows the correct answer for the set of selected input vectors.

Unsupervised learning is used when, for a given input, the exact numerical output a network should produce is unknown. Assume, for example, that some points in two-dimensional space are to be classified into three clusters. For this task we can use a classifier network with three output lines, one for each class. Each of the three computing units at the output must specialize by firing only for inputs corresponding to elements of each cluster. If one unit fires, the others must keep silent. In this case we do not know a priori which unit is going to specialize on which cluster. Generally we do not even know how many well-defined clusters are present. Since no "teacher" is available, the network must organize itself in order to be able to associate cluster with units.

Supervised learning is further divided into methods which use reinforcement or error correction.

Reinforcement learning is used when after each presentation of an input-output example we only know whether the network produces the desired result or not. The weights are updated based on this information so that only the input vector can be used for weight correction.

In learning with error correction, the magnitude of the error, together with the input vector, determines the magnitude of corrections to the weights, and in many cases we try to eliminate the error in a single correction step.

1.7 Back Propagation

As the algorithm's name implies, the error (and therefore the learning) propagate backward from the output nodes to the inner nodes. So technically speaking, back propagation is used to calculate the gradient of the network with respect to the network's modifiable weights.

The back propagation algorithm looks for the minimum of the error function in weight space using the method of gradient descent [8]. The combination of weights which minimizes the error function is considered to be a solution of the learning problem. Since this method required computation of the gradient of the error function at each iteration step, we must guarantee the continuity and differentiability of the error function. Obviously we have to use a kind of activation function other than the step function used in perceptrons, because the composite function produced by interconnected perceptrons is discontinuous, and therefore the error function too.

Because the network is equivalent to a complex chain of function composition, we expect the chain rule of differential calculus to play a major role in finding the gradient of the function. We take account of this fact by giving the nodes of the network a composite structure. Each node now consists of a left and right side.

2. METHODOLOGY

2.1 Steps of the Back Propagation Algorithm

We can now formulate the complete back propagation algorithm and prove by induction that it works in arbitrary feed-forward networks with differentiable activation functions at the nodes. We assume that we are dealing with a network with a single input and a single output unit.

2.2 Feed-forward

The input x is fed into the network. The primitive functions at the nodes and their derivatives are evaluated at each node. The derivatives are stored.

2.3 Back propagation

The constant 1 is fed into the output unit and network is run backwards. Incoming information to a node is added and the result is multiplied by the value stored in the left part of the unit. The result is transmitted to the left of the unit. The result collected at the input unit is the derivative of the network function with respect to x.

Hence by using the above mentioned Artificial Neural Network (trained by back by back propagation); we have formulated the model for the given objective function.

2.4 Algorithm

1. Draw 20*10 matrix of various inputs that may be financial ratio or any other parameter of the companies belonging to the sector under consideration of the listed companies. In the proposed model we have utilized the following inputs for calculation :

1. Share Capital

2. Reserve & Surplus

3. Secured Loans

4. Unsecured Loans

5. Fixed Assets

6. Current Assets

7. Current Liabilities

8. Pre Operating Expenses

9. Sales

10. Closing Stock

11. Opening Stock

12. Purchases

13. Freight Inwards

14. Gross Profit

15. Other Income

16. Administrative expenses

17. Depreciation

18. Profit before tax

19. Provision for Tax

20. Fringe Benefit Tax

Net Profit (output parameter)

Net Profit has been selected as the output of the network.

2. Apply Correlation for selection of independent variable. To know that inputs taken are how much dependent on each other and to also find the least dependent i.e. independent variable. Correlation gives the dependency as correlation factor r of x and y inputs.

3. Do the selection of the most independent variable which will be further taken as inputs for the neural network. Procedure of selection of very important variables is similar to Vogel's approximation method. It is an effective application of the best cell method along with some tie-breaking features.

When we apply the techniques of VAM on the matrix, if more than one "cell" is competing for allocation, tie-breakers are used in the following sequence:

Make the allocation on the cell with the smallest one.

In the case of tie in (1), make allocation to the cell with the largest possible allocation.

In the case of tie in (2), make allocation to that cell with first occurrence.

In an industry there are several factors which influence the future prospect of the company. But at a random it is difficult to say which are the most important factors which should be considered are as such. So we again need a mathematical tool for the taking out the taking out the factors we will plot all the factors in a matrix with their recent values and by using techniques of VAM we will be eliminating the factors. We will continue eliminating factors until we get the required number of factors for calculation. The techniques of VAM are mathematical tool which will help us to justify the selection of factors.

4. Set the network parameters. Take the inputs as selected, apply weights randomly and take an output parameter.

5. Input the data from file and set input and output parameter.

6. Find out the minimum and maximum of each row and column.

7. Compute scale value. It is calculated to bring all values between 0 and 1.

8. Assign threshold and weights using random numbers. Threshold value must not be greater than 0.5.

9. Compute prevailing value of each cell using prevailing value of previous cell.

10. Repeat this step for all inputs.

11. Compute the difference between original value and calculated value. It will be our error. It may be positive or negative. Check whether it is in the permissible limits or not.

12. If the error is in acceptable limits then calculate threshold and weights for each cell and set the network parameters.

Else adjust the threshold and weight by back propagation technique in which we find a learning factor through error and which helps us in assigning new weights and values and this step is carried out till we achieves a permissible error. And after that set the network parameters and network is ready to use.

2.5 Univariate Analysis

Univariate analysis is one of the methods for analyzing data on a single variable at a time. Univariate analysis explores each variable in the data set, separately.

Steps involved in Univariate Analysis

1. Determining the most influential input parameter

Step 1

The value of 1st input parameter is increased by 1 keeping other parameter constant i.e. maintaining their original values and the output is calculated.

Step 2

The value of 2nd input parameter is increased by 1 keeping other parameter constant i.e. maintaining their original values and the output is calculated.

Step 3

The value of 3rd input parameter is increased by 1 keeping other parameter constant i.e. maintaining their original values and the output is calculated.

Step 4

The value of 4th input parameter is increased by 1 keeping other parameter constant i.e. maintaining their original values and the output is calculated.

Step 5

The output values which is calculated in the previous steps (step1-4) is then compared and the input parameter giving maximum output value is considered as the most influential parameter.

2. Determining the length of influence

Once the most influence parameter is identified then same process is repeated till the other 3 input parameter does not give the maximum output as compared to the influential parameter(in this step the value of most influential parameter is increased by 1 every time for the further process).

3. CONCLUSION

The use of ANN gives an edge over the other techniques.

The proposed model is not based on assumption but based on supervised learning using back propagation algorithm, thereby enhancing its applicability and accuracy.

The proposed model will help each and every organization to understand the sensitivity, criticality and limits of variation of each parameter, which will help the management to earn more profits and reduce their losses.

4. REFERENCES

[1] Pradeepta Kumar Sarangi, Prashant Sarangi, 'Short term load forecasting using Neural Network Technology' The IUP Journal of Computer sciences, Vol.IV No.2, April 2010, pg.22

[2] M Balakrishnan, K Meena 'ANN model for coconut yield prediction' The IUP Journal of Computer sciences, Vol.IV No.1, Jan 2010, pg.28

[3] Seref Oruc, 'Neural network model for temperature sensivity of emulsified asphalt mixtures' Indian Journal of Engineering & Materials Sciences, Vol.17, Dec. 2010, pg.438.

[4] Prof. S K Shah, J Gadit, 'Genetically optimized ANN for nonlinear system identification' Journal of the Institution of Engineers(India), Vol.90, March18, 2010, pg.59

[5] D T Pham and S J Oh 'identification of plant inverse dynamics using neural networks in artificial intelligence in engineering', vol 13, 1999 pg. 309.

[6] Simon Haykin, "Neutral Networks". A Comprehensive Foundation, 2nd Edition, Pearson Education.

[7] Mohammad H. Hassoun, "Fundamental of Artificial Neutral Networks", Prentice Hall of India.

[8] B. Yegnanararayana, "Artificial Neutral Networks", Prentice Hall of India