

Artificial Neural Network Based Model to Estimate Profit for SMEs

Ajay Batra¹, Dr. Hullash Chauhan²

¹Mechanical Engineering Department, YITM, Rajnandgaon, CSVTU, Chhattisgarh, India

²Dean, Bharti University, Durg, Chhattisgarh, India

¹ajaybatra2106@gmail.com, ²hullashchauhan@gmail.com

Abstract

The behavior of any company is highly complicated. This work aims to evaluate of financial performance of a sector-specific business. Numerous financial indicators can be considered, but they often bear only a weak relationship to output performance. Several efforts have been made by researchers to establish stronger input–output relationships. Neural networks provide a powerful technique to capture such relationships, as they can handle non-linearities with ease. This study employs an Artificial Neural Network (ANN)–based model to estimate profit using four independent parameters for small and medium enterprises (SMEs). The output obtained shows a lower error rate compared with traditional Regression Analysis. The established relationship offers deeper insights into the intricate behavior of the sector, allowing for a more precise analysis of the criticality of parameters under consideration. The findings may be useful to companies, board members, shareholders, 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 processed to achieve desired outputs over a period of time. There exists a relationship between these input and output variables. Management decision-making can be significantly enhanced if a model is developed that not only establishes these relationships but also identifies the criticality of the variables involved. Addressing this need, the present work proposes a model for the given problem. At present, SMEs have been considered for the study.

1.1. Problem Definition & Objective Function

“To design a model that can establish a relationship between the input and output variables of an industrial sector (SME) in order to predict the output & analyze the influence of different input parameters, thereby establishing their critically and the degree 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 interrelationships among the variables, which are studied and corrected to a considerable extent to aid in predicting and analyzing industry performance. The model uses artificial neural networks to analyze and predict future outcomes.

The model is expected to benefit both existing companies and new entrants in the sector. An artificial neural network is a system inspired by the functioning of biological neural networks. It is an adaptive, typically nonlinear system that learns to perform a mapping function from input to output data.

Correlation analysis is employed to justify the selection of input parameters for the ANN. These parameters are iteratively adjusted during the training phase, after which the ANN parameters are fixed and the system is deployed to solve the target problem.

Additionally, univariate analysis is performed to identify the most influential input parameter affecting the output.

1.2. Available Methods and Selection of Best Alternative

1. Regression Analysis

2. Artificial Neural Network (ANN) integrated with correlation

1.2.1. Reasons for Selecting ANN as the Best Alternative

- Regression analysis is often subject to considerable manipulation.
- The forecasted results obtained through regression analysis showed significantly higher errors (difference between actual and predicted values)
- A neural network can perform tasks that linear programs cannot.
- Due to their parallel nature, neural networks can continue functioning even if some elements fail.
- Neural networks are capable of self-learning and do not require reprogramming.
- The Back Propagation method has been found to be a strong forecasting tool, offering advantages over regression analysis [1].

2. Literature Review

2.1. Artificial Neural Network (ANN)

Artificial neural networks are computer systems that can learn from the features of nervous system, derive new information using the new information learned, and work similarly to decision making structure. ANN has emerged as a result of mathematical modeling of the learning process by taking the human brain as an example [2]. As shown in Figure2.1, the neural network processes a number of inputs to obtain a number of outputs. 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 [3].

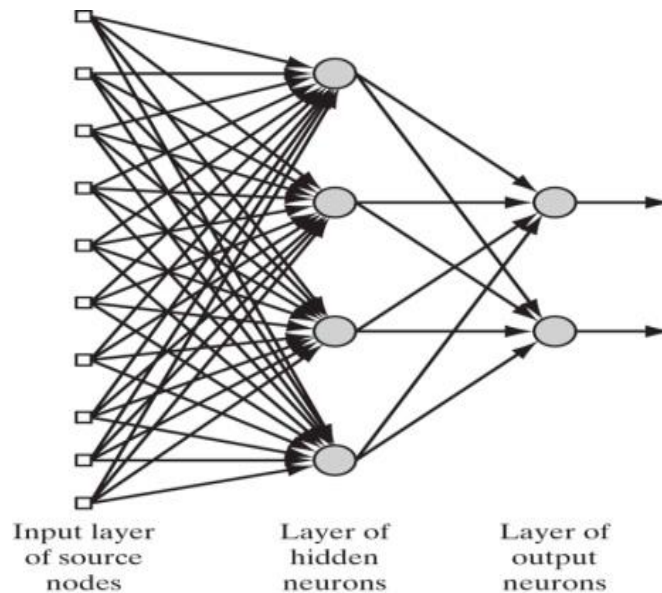


Figure 2.1 Artificial Neural Networks

2.2. ANN along with “Correlation”

Correlation is the most widely used method of measuring the degree of relationship between two variables [4]. It is a statistical technique that indicates whether, and how strongly, pairs of variables are related. The result of this computation is the correlation coefficient (r).

The calculation of the correlation coefficient for two variables, say x_i and x_j , is given by

$$r(x_i, x_j) = \frac{n \sum x_i x_j - \sum x_i \sum x_j}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum x_j^2 - (\sum x_j)^2}} \quad (2.1)$$

Where, n is the number of pairs of data.

In industrial analysis, many factors are responsible for performance. However, an important question arises: Are these factors independent, or do they influence one another? Correlation analysis helps answer this by quantifying the degree of dependence between variables. The 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 [4].

Selecting the least dependent variables (i.e., those with low correlation) is crucial for ANN modeling, as this ensures that the inputs provide unique information.

2.3. Back Propagation

Back propagation (BP) is the most popular training algorithm for multilayer Neural Networks [5]. Back propagation algorithm is popular for its simplicity of implementation and its ability to quickly generate networks that have the capacity to generalize [6]. As the name suggests, errors propagate backward from the output layer to the hidden layers, allowing the model to adjust weights.

The BP algorithm seeks to minimize the error function in weight space using gradient descent. The set of weights that minimizes the error function is considered the solution to the learning problem. Since BP requires the error function to be continuous and differentiable, it is technically used to compute the gradient of the network with respect to its modifiable weights. Neurons in the hidden layer are assumed to be sigmoidal and those in output layer are assumed as linear [7].

Multi Layer Perceptron (MLP) is a popular form of ANN. Moreover BP is a well-known gradient-based approach for training MLP [8]. It is very successfully used in many applications in various domains such as prediction, function approximation and classification. Mean Square Error (MSE) has been considered to obtain the optimal NN model among all the NN models developed. The mean square error (MSE) criterion is given by:

$$MSE = \frac{1}{2} [\sum_p \sum_k |t_{kp} - O_{kp}|^2] \quad (2.2)$$

Where t_{kp} and O_{kp} are the true and observed outputs, respectively, for neurons k in the output layer when input vector x_p corresponding to the p^{th} training record is applied to the network. The root-mean square error (RMSE) also describes the average magnitude of the errors, given by

$$RMSE = MSE^{0.5} \quad (2.3)$$

Since the errors are squared before they are averaged, more weight is given to large errors [9].

The output from neuron j in a given layer (other than the input layer) is calculated as:

$$O_j = f(\sum_i W_{ij} O_i) \quad (2.4)$$

Where i indicate a neuron in the preceding layer and $f(x)$ is the activation function for neuron j . The activation function is often a sigmoid function of the form:

$$f(x) = \frac{1}{1+e^{-x}} \quad (2.5)$$

With the gradient descent approach to error minimization, weights are changed in proportional to the error gradient, “i.e.”,

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (2.6)$$

Where η is a constant that determines the learning rate. To improve convergence characteristics, a weight change is given by (Park, Marks, Atlas, & Damborg, 1991):

$$\Delta W_{ij}(n) = \eta \delta_j O_i + \alpha \Delta W_{ij}(n-1) \quad (2.7)$$

Where η is the learning rate, α is the momentum factor, and δ_j is the error signal for the destination neuron j . When neuron j is the output layer, δ_j is given by:

$$\delta_j = (t_j - O_j) O_j (1 - O_j). \quad (2.8)$$

When neuron j is in a hidden layer, δ_j is given by:

$$\delta_j = O_j (1 - O_j) \sum_k \delta_k W_{jk}, \quad (2.9)$$

Where k indicates neurons in the succeeding layer next to that containing neuron j . [10].

The learning rate and the momentum factor influence the speed and stability of network training [11].

The process continues until the error criterion on the training set is reduced below a specified limit.

2.4. 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. In this method we change only one variable at a time and seek to produce a sequence of improved approximation to the minimum point. The Univariate method is very simple and can be implemented easily. However, it will not converge rapidly to the optimum solution. Hence it will be better to stop the computation at some point near to the optimum point rather than trying to find precise optimum point [12].

3. Methodology

3.1. 1 Algorithm

Step 1. Draw 20*12 (m*n) matrix of datasets.

Step 2. Here, 20 parameters are considered (financial ratios or other indicators), taken from the audited financial accounts of several companies across *n* financial years.

In the proposed model, the following inputs are utilized:

- Share Capital
- Reserve & Surplus
- Secured Loans
- Unsecured Loans
- Fixed Assets
- Current Assets
- Current Liabilities
- Pre Operating Expenses
- Sales
- Closing Stock
- Opening Stock
- Purchases
- Freight Inwards
- Gross Profit
- Other Income
- Administrative expenses
- Depreciation
- Tax
- GST
- Net Profit(selected as the output of the network).

Step 3. Apply correlation analysis for selection of independent variables. This step identifies the degree of dependency among input variables and helps in selecting the most independent (least correlated) variables for the ANN.

Step 4. Do the selection of the most independent variable which will be further taken as inputs for the neural network.

Step 5. Set the network parameters.

Step 6. Input the data from file and set input and output parameter.

Step 7. Find out the minimum and maximum of each row and column.

Step 8. Normalize the data by computing the scaled values.

$$\text{Scaled value(SV)} = \frac{\text{Original value} - \text{min value}}{\text{Max value} - \text{min value}}$$

This transformation brings all values between 0 and 1.

Step 9. Assign threshold and weights using random numbers.

Step 10. Compute the forward pass:

Calculate the prevailing value of each neuron based on the preceding layer's outputs.

Step 11. Repeat this process for all input parameters.

Step 12. Compute the error (MSE/RSME). Check whether it falls within acceptable limits.

If the error is in acceptable limits then finalize the thresholds and weights, and set the network parameters.

If the error is not acceptable, then adjust the thresholds and weights using the back propagation technique. This process iteratively updates the weights until the permissible error is achieved. Once convergence is reached, the network is considered trained and ready for use.

3.1.2. Application of Univariate Analysis

Suppose there are four selected input parameters.

- Keep any three inputs constant and increase the value of the fourth input. Record the change in the output.
- Repeat this procedure for each input parameter.
- The parameter that produces the largest change in the output is identified as the most influential parameter.

3.2. Deciding Network Topology

Using more number of layers will help the network learn faster. Place 2 neurons in each layer and train the network and test for the performance. Increase the number of neurons in a layer and train the network again till satisfactory performance. This systematic procedure helps to obtain an optimal NN architecture [13].

3.3. Steps involved in Univariate Analysis

Determining the most influential input parameter

Step 1: Increase the value of the first input parameter by 1 unit while keeping all other parameters constant at their original values. Record the resulting output.

Step 2: Increase the value of the second input parameter by 1 unit, keeping all others constant, and record the output.

Step 3: Repeat the process for the third input parameter.

Step 4: Repeat the process for the fourth input parameter.

Step 5: Compare the outputs obtained in Steps 1–4. The input parameter that produces the maximum change in the output is identified as the most influential parameter.

4. Results & Conclusion

Most of the available prediction techniques are generally based on assumptions. In contrast, the proposed model relies on supervised learning through the back propagation algorithm, which enhances both its applicability and accuracy. The model learns from existing and historical data, adjusts itself dynamically, and applies the necessary modifications, thereby enabling highly accurate predictions of the target parameter.

The results demonstrate that the ANN-based model produces lower errors compared to regression analysis, confirming its superior predictive capability. Furthermore, the model not only predicts outputs but also provides insights into the sensitivity, criticality, and permissible variation limits of each parameter.

Such information is of practical value to organizations, as it can assist management in identifying key financial drivers, improving decision-making and ultimately maximizing profits while minimizing potential losses.

5. Scope of Future Work

- The proposed model provides a strong foundation, yet several avenues remain open for further enhancement:
- The model can be fine-tuned by establishing a more detailed sensitivity relationship between the input and output variables.
- Predictive capability may be improved by analyzing additional financial and non-financial aspects of the SME sector.
- An extension of the current framework could involve grouping certain inputs into clusters, thereby incorporating mixed or composite input variables.
- The simplified ANN approach may be integrated with fuzzy systems or rule-based systems, resulting in hybrid neuro-fuzzy models.
- Alternative activation functions for hidden layers can be explored to test whether they yield superior performance compared with the sigmoid function.
- Advanced stochastic optimization techniques such as simulated annealing and genetic algorithms could be applied to overcome the issue of local minima.
- Future studies could leverage the variety of learning algorithms available in MATLAB (e.g., Gradient Descent with Momentum, Levenberg–Marquardt) to evaluate and compare performance.

References

1. Ying Deng , et al., “New methods based on Back Propagation (BP) and Radial Basis Function (RBF) Artificial Neural Networks (anns) for predicting the occurrence of haloketones in tap water”, Elsevier, Journal of Science of The Total Environment, Volume 772, 10 June 2021, 145534, ISSN 0048-9697.
2. Badr Malek, et al., “Detection of Heart arrhythmia on Electrocardiogram using Artificial Neural Networks” Journal of Computational Intelligence And Neuroscience, Special Issue , Recent Advances in Multimodal Environment for Biomedical Diagnosis and Computational Analysis, ,volume 2022, 05 August 2022.
3. Jianali Feng, Shengnan Lu, “Performance analysis of various activation functions in Artificial Neural Networks” , IOP Conf. Series Journal of physics: 1237 (2019) 022030.
4. Kothari C.R., (2019), “Research Methodology”, Methods and Techniques, 4th Edition, New Age International Private Limited.
5. Kassaymeh Sofian, et al., “Back propagation Neural Network Optimization and Software Defect Estimation Modeling using a hybrid Salp Swarm Optimizer- Based Simulated Annealing Algorithm”. Science Direct, Journal of Knowledge- Based Systems, volume 244, 23 may 2022, 108511, ISSN 0950-7051.
6. Satish Kumar, (2017). “Neural Networks”. A Classroom Approach, 2nd Edition, TMH Education Private Limited.
7. Arya L D, Koshti A, “Identification of static load models using ANN” Journal of the Institution of Engineers(India), Vol.88, March28, 2008, pg.28-31.



8. Agahian S., Akan T., “ Battle royale optimizer for training multi-layer perceptron”, Springer , Evolving Systems 13, 563-575(2022).
9. Nowrin T., Kwon T.J., “Forecasting shorts- term road surface temperatures considering forecasting horizon and geographical attributes – an ANN based approach” journal of cold regions science and technology, volume 202, October 2022.
10. Varaprasad B.J.S., Viswanadh G.K., “Artificial Neural Network model for estimation of deposits formation in PVC pipes”, Indian Journal of Neural Network Research”, Vol. 2, No. 1, Jan-June 2012, pg.1-5.
11. Mohamad Hassoun H., (2007), “Fundamentals of Artificial Neutral Networks”, Prentice Hall of India.
12. Rao S.S., (2006) “Engineering Optimization”. Theory and Practice, 3rd Edition, New Age International Private Limited.
13. Muthuramalingam A., Himavathi S., Srinivasan A., “Neural Network Implementation Using FPGA: Issue and Application”, International Journal of Information and Communication Engineering Vol. 4, June 2008, pg.396-402.