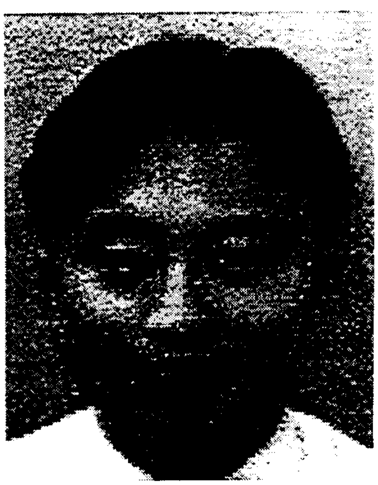
A REVOLUTIONARY APPROACH TO COST PREDICTION USING THE ARTIFICIAL NEURAL MODEL

Deepak Burman, MS; Value Engineering Manager - Applied Materials, Inc.

Deepak Burman, MS is a Value Engineering manager for Applied Materials, Inc. He is responsible for Design to Cost and Cost Reduction activities for the spares and consumable division. Mr. Burman has over 20 years of experience in design, materials, process and Value Engineering in the hi-tech industry. He received his BS in Mechanical Engineering and Masters of Science in Material Science from Polytechnic Institute of New York, New York.

Kae Huang, BS; Value Engineer - Applied Materials, Inc.



Kae Huang received a Bachelor's of Science degree in Industrial Engineering from Purdue University. Graduating with Honors and with Distinction, he assisted in the research of machine learning, specifically Feature Subset Selection Using Nonlinear Correlation Analysis (FSSNCA). Only recently has he begun practicing Value Engineering and attended the SAVE International approved Value Methodology Workshop. With the mix in his background, he is bringing with him a new technique for VE cost analysis - the artificial neural network model.

ABSTRACT

This paper introduces the artificial neural network concepts and describes how a model can be developed. It follows with a discussion of its application in manufacturing cost estimation and Value Engineering. Neural models can be extremely accurate and able to provide with results in a matter of minutes. This paper ends with a presentation of an actual model that was developed to estimate the cost of O-ring seals.

INTRODUCTION

Originally artificial neural networks were developed to model the nervous system operations and reproduce human information processing of speech, vision, touch and others. Later, it was further developed for function approximation, pattern recognition and other business applications. These include: analysis of market research data, manufacturing process control, forecasting, flight control, prediction of runs scored in a baseball game, and many others.

HOW IT WORKS

Artificial neural network, as its name implies, processes information in a way similar to how the human brain works. Indeed, artificial neural network can learn from experience. Artificial neural network is an interconnected system of "neurons," or processing elements, PEs. It can process large amount of data in parallel due to the connection pattern between the PEs. The structure of an artificial neural network is somewhat as shown:

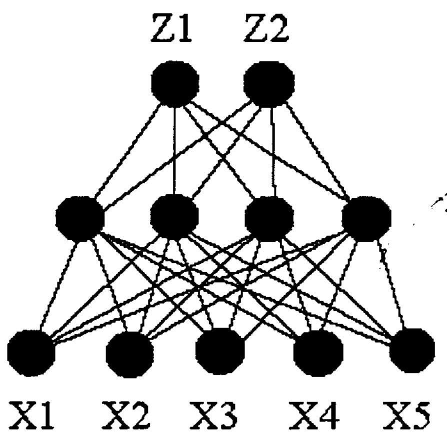


Figure 1

As indicated above, PEs are structured into layers. The bottom layer is the input layer, or the input variables or fields. In this case there are 5 of them. The middle layer is the hidden layer, where most of the work is done. This layer has full connection with the inputs, therefore it learns about the relationship between the input variables. Finally, the upper layer is the output layer, or the output variable or result that we are searching for. Usually there is only one output node, but sometimes there might be more. As the artificial neural network is trained, the strength of each connection increases or decreases based on the "weight" it has on the final output.

WHY NEURAL NETWORKS

This is the information age. It is critical that business organizations have the capability to better manage data and information and utilize it to make better business decisions. Today, artificial neural network seems to be one of the most sophisticated and best methods in learning from historical data. Below are some of the key strengths of artificial neural network:

- ability to provide results and extract key findings much faster than many other tools
- ability to determine non-linear relationship between the input variable(s) and the output variable(s)
- ability to search for all possible interrelationships among input variable(s)

DEVELOPING A NEURAL MODEL

The basic steps towards the development of an artificial neural model consist of:

1) Defining Objective
2) Identifying Variables
3) Collecting Data
4) Transforming Data
5) Selecting Train, Test, and/or Validation Sets
6) Network training
7) Model Verification

The first three steps will have to be developed solely by the individual. Most of the commercially available neural network software will complete the next three steps. Finally, step 7 will be completed with the combined effort of the user and the model. In the next section, a case study will be used to get into more details. A key point to remember about neural network is that the process alone of analyzing the data can be invaluable. Often times key findings are obtained in this process, before the neural model is even trained.

NEURAL NETWORK IN COST ESTIMATION AND VALUE ENGINEERING

Most of the research and development of artificial neural networks is for applications such as data mining, forecasting, system modeling, optimization, and others. Clearly, neural models should also be able to predict manufacturing costs. For example, to determine the cost of a vehicle, we would need input variables such as: the number of cylinders, body style, whether there is an Anti-lock Break System, where it was made, and so on.

Whenever there is a substantial amount of historical data for a particular manufactured product and it takes at least hours to determine the cost using the traditional DFM software, it would be worth a while developing a neural network model. If validation results indicate that the model is accurate, cost estimating that same product in the future will only take a few minutes. If validation shows that the model is not accurate enough for practical purposes, it can at least be used as a quick estimate report.

But most importantly, as previously mentioned, the process alone of analyzing the data for neural networks can be invaluable. Key findings, thus value analysis, can be made during this process, providing a VE with ideas to improve the design or manufacturing process of the product. For example, while comparing the data of two fabricated plates, one may find that they are almost identical, only difference being one of them has an extra pocket which must be masked before plating. With that single difference, its cost is twice as much as the other plate. As a Value Engineer, it might be a good idea to somehow eliminate that feature and reduce the cost by 50%.

A CASE STUDY - COST OF O-RING SEALS

O-Rings are in fact a significant part of the overall cost and also one of the largest consumable in Applied Materials systems. To keep the confidentiality of our cost information, the data has been altered without sacrificing the purpose of this exercise. The objective is to develop a quick method to estimate the cost of an O-ring. The possible input variables identified were the outer diameter, grove diameter, order quantity and compound specification. The data size was 42 part numbers, (though only 37 were exposed to the model) and the data was obtained from the individual part description and definition forms.

The demo software used to model the cost of the O-Ring is called NeuralWorks® *Predict™*, developed by Aspen Technology, Inc. It works within a Microsoft Excel spreadsheet, with Predict included in the menu bar. NeuralWorks Predict has the capability to transform data, select the train, test and validation data sets, and of course create the model. In addition, it provides a statistical summary report. NeuralWorks Predict does indeed do most of the work after step three of the neural network model development mentioned earlier.

After loading the data into Microsoft Excel, create a new file for the new model and the wizard will direct the user throughout the process. The software will prompt the user to identify the location of the input and output variables and the total size of the data set. It will also prompt for other parameters such as the Problem Type, Noise Level, Data Transformation, Variable Selection and Network Search as shown in Figure 2.

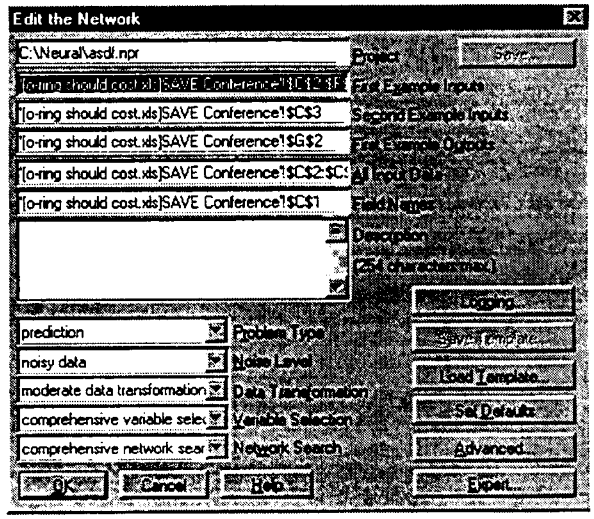


Figure 2

The next step is to just let the model run. A screen as shown in Figure 3 will appear - this is the period when the neural network is learning and creating the model.

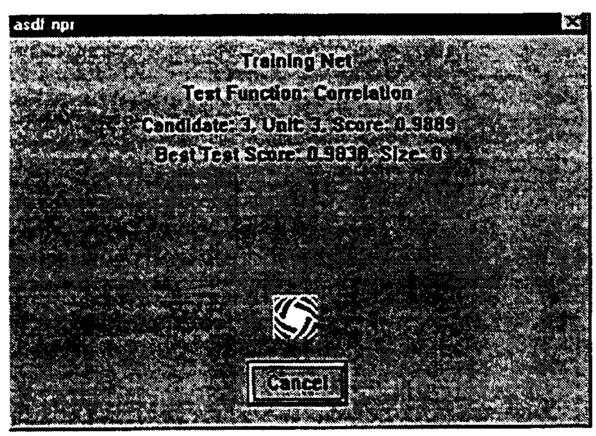


Figure 3

RESULTS - O-RING MODEL

For this case study, it only took around 20 seconds to run the model. Figure 4 plots the actual, neural and linear regression values for each O-ring part number. According to this chart, both the neural and linear regression plots seemed to predict the value fairly well.

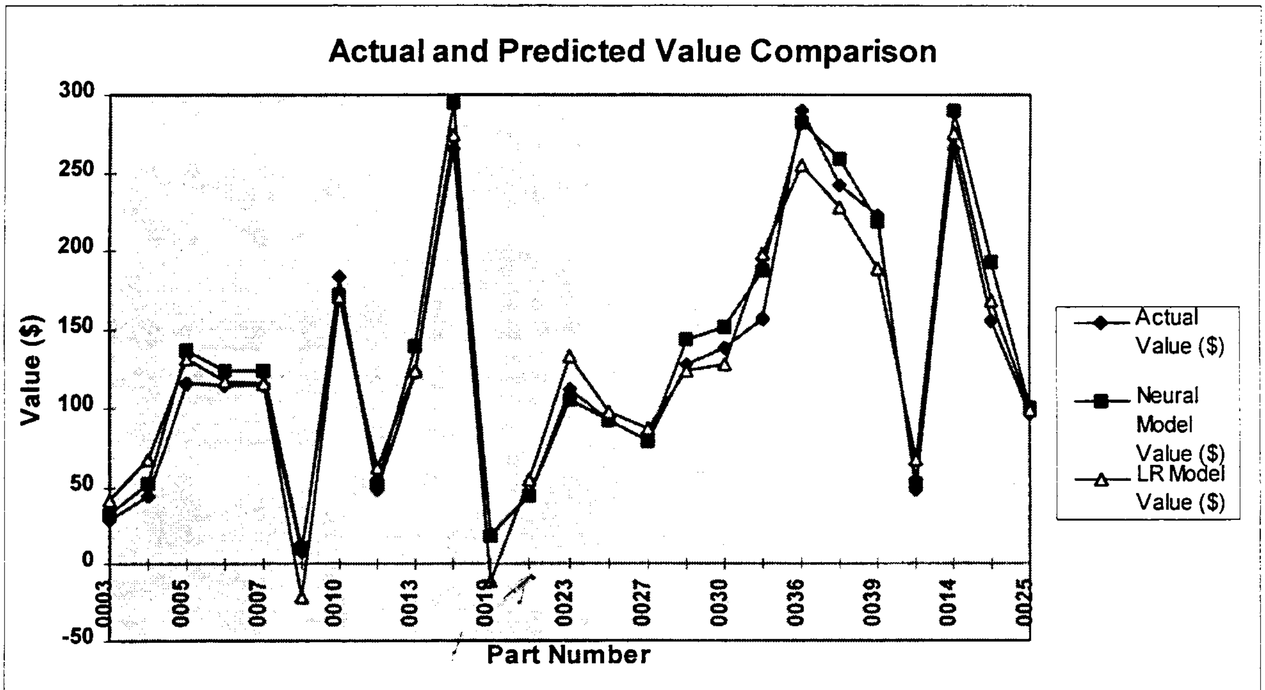


Figure 4

To further verify results a statistical analysis as shown in Figure 5 was originated. Although 42 records were available, Figure 5 only identifies 25 (21 exposed to the neural model and 4 for testing after training) because other records contained missing fields and the linear regression model does not accept incomplete records. The first six columns on top of Figure 5 contain the data that was originated to create the model. Next 4 columns are the neural model results and the last four are the linear regression calculations. The *average absolute delta %* of the neural model cost estimate in relation to the actual historical value for the parts shown in Figure 4 was 10%. On the other hand, it was 35% for the linear regression model. The *R value* obtained from the linear regression model was 0.94 (indicated in the Linear Regression Model Table), whereas for the neural model, it ranged from 0.972 (*Training* data set) to 0.995 (*Manual* data set). One of the main reasons for the lower *R value* for the training data is due to outliers, mainly part number *0008*. The delta % for this particular part number was 32 from the neural model; the linear regression model was way off with a delta of 375%. The *R value* shows positive results for the *Manual* data set of 4 part numbers. This is the set that was absent during the training process; nevertheless, the model was able to predict it with great accuracy. It can be concluded that the model was not overtrained or overfit. Another positive result came from the *Accuracy (20%)*, where for the test and manual data set, it came out to be 100%!!. On the other hand, *Confidence Interval (95%)* did not come out as strong with values in the 30s to 50s.