Outlines the differences and similarities between the two techniques.

Neural Networks and Statistical Techniques in Marketing Research:

A Conceptual Comparison

V. Venugopal and W. Baets

Marketing Intelligence & Planning, Vol. 12 No. 7, 1994, pp. 30-38

Introduction

Recently, there has been considerable interest in the development of artificial neural networks (ANNs) for solving a wide range of problems. Artificial neural networks are distributed information-processing systems composed of many simple computational elements interacting across weighted connections. Inspired by the architecture of the human brain, ANNs exhibit certain features such as the ability to learn complex patterns of information and generalize the learned information.

The authors acknowledge the financial support given by the European Foundation for Management Development for this research.

Computer scientists have exploited these features and achieved a breakthrough in vision- and speech-processing technology. ANNs are no longer considered to be the subject of computer science alone. They are highly interdisciplinary and have attracted the attention of researchers from different disciplines. Their utilities have been demonstrated in wide areas ranging from engineering to management problems (Nelson and Illingworth, 1991). The use of ANNs in business practice is gathering momentum day by day.

In particular, the ability of neural networks to identify patterns in the data could be utilized in market research, especially in areas which were once reserved for multivariate statistical analysis. For this reason, neural networks are often considered to be statistical methods (White, 1989).

If neural networks are considered merely as statistical techniques, then why has so much importance been given to them during recent times? What are the special features they have over traditional statistical techniques? These are the questions that have been raised time and again among researchers. For a better understanding and utilization of ANNs, it is essential to know conceptually the difference between neural networks and conventional statistical techniques. As such, market researchers and managers who are not aware of the conceptual difference between these two methods cannot use this technology effectively. This article is a step in the right direction.

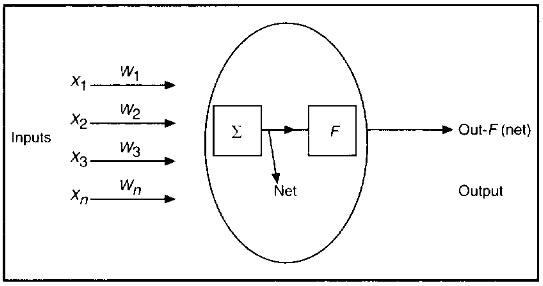
The article emphasizes the conceptual differences and similarities between neural networks and statistical techniques in handling certain problems of business research. It is meant for market researchers and managers who are looking for new tools to support their decision making.

The article is organized into four major sections. First, an overview on neural networks is presented. This is followed by a discussion of forecasting problems and regression analysis/neural networks as tools for solving such problems. The next two sections outline the classification and grouping problem and give an overview of statistical and neural network approaches for solving such problems. Finally, ANN's properties, advantages and disadvantages over conventional statistical tools such as regression analysis, cluster analysis and discriminant analysis, are emphasized.

Overview of Artificial Neural Networks

Artificial neural networks attempt to model the architecture of biological neural systems. Biological neural networks are made of simple, tightly interconnected processing elements called neurons. The interconnections are made by the outgoing branches, the

Figure 1. *Artificial Neuron*



"axons", which again form several connections ("synapses") with the other neurons. When a neuron receives a number of stimuli, and when the sum of the received stimuli exceeds a certain threshold value, it will fire and transmit the stimulus to adjacent neurons. The aim of ANNs is to extract concepts from the biological networks with which powerful new computational methodologies can be developed.

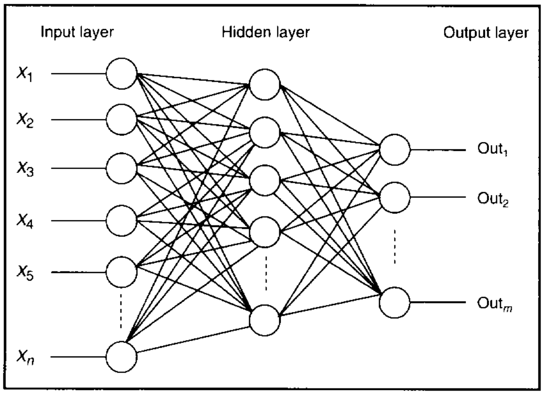
Artificial neural networks consist of many non-linear computational elements called nodes. The nodes are densely interconnected through directed links. Nodes take one or more input values, combine them into a single value, then transform them into an output value. Figure 1 illustrates a node that implements the macroscopical idea of biological neurons.

In Figure 1 a set of inputs labelled $X_1, X_2, \ldots X_n$ is sent to a node. Each input is multiplied by the weights of the interconnections $W_1, W_2, \ldots W_n$ before it is applied to the summation block. Each weight corresponds to the strength of a synaptic connection. The summation block adds all the weighted inputs algebraically, producing an output denoted as NET. The block labelled F accepts the NET output. If the NET output exceeds the threshold level, the OUT node is said to be activated.

The power of neural computing comes from connecting artificial neurons to artificial neural networks. The simplest network is a group of neurons arranged in a layer. Multilayer networks may be formed by simply cascading a group of single layers. Figure 2 shows a three-layer neural network: an input layer, an output layer, and between the two a so-called hidden layer. The nodes of different layers are densely interconnected through directed links. The nodes at the input layer receive the signals (values of the input variables) and propagate concurrently through the network, layer by layer.

The numbers of layers and neurons, and the weights to be attached to the connections from neuron to neuron, can be decided in such a way that they give the best possible fit to a set of data. Different types of neural network models have been developed in the literature.

Figure 2. *Schematic Representation of Three-layer Neural Networks*



ANN models are characterized by their properties: namely, the structure of the network (topology), how and what the network computes (computational property) and how and what the network learns to compute (learning or training property). Learning is the process in which a set of input values is presented sequentially to the input of the networks and the network weights are adjusted such that similar inputs give the same output. Learning strategies are categorized as supervised and unsupervised.

Supervised learning requires the pairing of each input value with a target value representing the desired output and a teacher who provides error information. In unsupervised learning, the training set consists of input vectors only. The output is determined by the network during the course of training. The unsupervised learning procedures construct internal models that capture regularities in their input values without receiving any additional information.

The massive number of processing elements makes neural computing faster than conventional computing. They are robust and fault tolerant owing to their parallelism. They are fault-tolerant in the sense that their performance does not degrade significantly even if one of the nodes fails. Also, based on current results, neural networks adapt themselves – that is adapt their structure and/or connection weights – to achieve a better performance. (For further details on neural networks see Nelson and Illingworth (1991).)

Forecasting Problem

Often, marketing managers have to make decisions with the knowledge of what will happen in the future. For example, market planners often need to forecast sales over the planning horizon. These forecasts are needed to make decisions on acquiring raw materials, managing labour and scheduling production. Retailers' sales forecasts are needed, especially when the items are stocked in a number of stores/locations to meet the local demand as it occurs. The retail sales forecastings are essential for efficient management of inventory at local stores so as to meet the demand (Thall, 1992). They are the basis of regional distribution and replenishment plans. Using databases on the size of trade area, competition, sales and prices of different items, distribution of population and other demographic characteristics of several stores, retailers have been making sales forecasts with the help of statistical methods. The most commonly applied statistical method is multiple regression analysis (MRA). The latest exciting development in forecasting is the use of neural networks. In this section, we discuss the regression and neural network approaches to forecasting.

Multiple Regression Analysis

Multiple regression analysis is a method for quantifying the relationship between a dependent factor/variable/criterion and one or more independent factor(s)/variable(s)/criterion(a). In the case of a product that is sold to a final consumer, MRA calls for determining the quantitative relationship with the following explanatory variables for determining retail sales forecasting:

$$Y = f(X_1, X_2, X_3, X_4, X_5, \ldots)$$

where, Y = market demand for the product
X_1 = consumers' disposable income
X_2 = size of the population
X_3 = price of the product
X_4 = price of substitutes
X_5 = price of complementary products

To formulate the equation, the manager/decision maker should know a priori the form of the equation which the available information represents. Under normal circumstances, a priori knowledge of the form of the equation is difficult. Hence the manager must specify a family of equations (regression curves) and select the member of the family that best fits the data. The quality of fit is measured in terms of an error term. The least square method is normally used to fit the equation.

Difficulties in Regression

The main difficulty in using regression analysis is the requirement of a priori knowledge of the functional form. As above, to formulate the equation, the manager/decision maker should know a priori the form of the equation which the available information represents. Under normal circumstances, a priori knowledge of the form of the equation is difficult. Otherwise, managers can try several functional forms and finally choose the one that best fits the data. Even in this worst case, the managers face the problem of deciding which functional form to consider for the problem under consideration.

Often, market researchers/managers make simplifying assumptions of linearity in the data structure, which has the advantage that models can be built more easily. But linear models are extremely bad at picking up turning-points in the available information. As marketing managers always deal with sales and price data, the data series is bound to have turning-points, trend and non-linearity. Sometimes, the data series may be chaotic also (Thall, 1992).

Besides these issues, several methodological problems such as multicollinearity and heteroscedasticity are involved in the regression analysis. For instance, in the above example, inventory and operating expenses are correlated and lead to multicollinearity problems.

The conventional multiple regression analysis can deal with only one dependent variable at a time. In our example, suppose the managers are interested in predicting profit besides cash flow, then these two cases are to be treated separately. When multiple factors are to be predicted, statistical analysis such as canonical correlation can help. But it is onerous to interpret the results of such an analysis and the methodology does not lend itself readily to making predictions (Proctor, 1992).

Market researchers may wonder if there is a methodology that can take all this into account and serve as a tool for accurate sales forecasts. Neural networks do, indeed, avoid these particular problems and can be of use to marketing managers to forecast sales.

Neural Network Approach to Retail Sales Forecasting

The prediction of retail sales and the time taken to break even, in the above example, can be made using a neural network approach. A network can be constructed in which the number of nodes at the input layer is equal to the number of independent variables, and the number of nodes at the output layer is equal to the number dependent variables. The number of hidden layers and the nodes in each hidden layer can be selected arbitrarily.

Neural networks with at least one middle layer use the data to develop an internal representation of the relationship between the variables so that a priori assumptions about the underlying parameter distributions are not required. As a consequence, better results might be expected with neural networks when the relationship between the variables does not fit the assumed model. Also, it has been proved that a network with only one hidden layer is enough to approximate any continuous function.

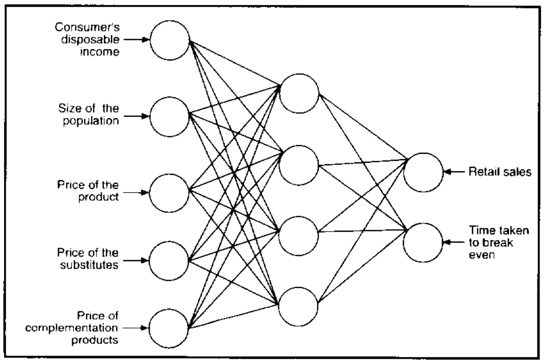
The middle layer nodes are often characterized as feature detectors that combine raw observations into higher order features, thus permitting the network to make reasonable generalizations. Too many nodes in the middle layer produce a neural network that memorizes the input data and lacks the ability to generalize. In most of the applications, the number of nodes in the hidden layer was taken to be at least 75 per cent of the number of input nodes.

For illustration purposes, we consider a 5-4-2 feed forward neural network model. The structure of the neural network is given in Figure 3. The network can be trained using the back propagation algorithm (Burke, 1991).

The neural network in Figure 3 consists of an input layer of five nodes, each of which represents X_1, X_2, X_3, X_4 and X_5 in our example; a hidden layer of four nodes; and an output layer of two nodes, which represents retailer sales forecasts and time taken to break even. Signals in the neural network feed forward from left to right. The network performs two operations, one at the hidden layer and one at the output layer.

Data on the independent factors, such as size of the population, price of the product and consumers' disposable income, and data on the dependent factors, such as sales and time taken to break even, form a signal. Initially, arbitrary values can be assigned to the weights of the network. Each case from a sample can be loaded on to the input layer of the network. The input nodes simply send these values to the hidden nodes. Each hidden node calculates the weighted sum of the inputs using the weights assigned to the connections. Each hidden node squashes the sum value down to a limited range and sends the result to all the output nodes. Each output node performs a similar calculation. The result of the calculation is taken as the value of the dependent variable, viz., sales and time taken to break even. Next, output nodes are given actual/observed value(s) of the dependent variables for that case.

Figure 3. *5-4-2 Neural Network for Retail Sales Forecasting*



Based on the difference between the computed value of the dependent variables and observed values of the dependent variables, each output node determines the direction in which each of its weights would have to move to reduce the error, as well as the amount of change that would be made, and this is propagated to a hidden node. The hidden nodes use these errors to determine in which direction and by how much they should change their weights, just as the output nodes did. This process, called training, is repeated over and over again to the network, enabling the network to adapt its weights so that the estimated cash flow and profit reflect its actual value. This process is measured by an error term, the difference between the estimated sales and the time taken to break even and its actual value summed over all signals. The goal of training is to minimize the error over all signals. After sufficient training, the network should be able to forecast. This can be tested with test data, composed of new facts not used in the training.

Managers do not even have to know the intricacies of neural networks and their working. With this idea, and with the help of neural network software tools such as Explorenet (HNC, 1990), managers can adopt a neural computing tool for their decision-making process. (For more details on software available on the market, see Jurik (1993).)

Neural Network Approach versus Regression Approach for Retail Sales Forecasting

The processing described in the previous section may look exactly like the regression method but the neural network approach forecasts retail sales by modelling the relationship between the independent factors and the dependent factors in a unique way. Neural network methods are non-parametric in the sense that functional form need not be specified a priori. Rather than relying on a prespecified functional form, neural networks build their own model by "learning", testing and modifying.

Neural networks adapt their weights as new input data become available, thus adjusting readily to a changing environment. A traditional regression method is not adaptive but typically processes all data once again together with the new data.

In neural networks, activation of each neuron is a separate linear combination. Thus neural networks, being the collection of neurons, define complex relationships. In general, we can say the relationships are built within neural networks. As a result, they handle non-linear/chaotic data series by deriving a suitable map between high-dimensional input pattern spaces and output.

Several methodological problems, such as multicollinearity and heteroscedasticity, are involved in the regression analysis. The conventional multiple regression equation can deal with only one dependent variable at a time. In our example, suppose the managers are interested in predicting the time required to break even, then these two cases are to be dealt with separately.

Unlike a regression equation, neural network can deal with more than one dependent factor at a time. The other feature of neural networks is that, if some of the data are missing, the networks can generalize across gaps by building up a model and interpreting it owing to its fault-tolerant nature. Neural networks perform well with missing or incomplete data, which is most difficult for regression. A single missing value in regression analysis calls for dropping the entire observation or dropping the variable from all observations.

There is no need to consider the statistical significance of the model parameters in neural networks. Testing the network is the same as cross-validation in the regression.

Researchers have tested the performance of ANN in forecasting. De Groot and Wurtz (1991) have compared neural networks with standard non-linear models and concluded that neural networks are the best when the data exhibit non-linear characteristics. Hruschka (1993) has compared the neural network model with an econometric model of market response. His results indicated that neural networks lead to better model fits compared with econometric models. White (1989) has brought out the similarity between the back-propagation method and stochastic approximation procedures. The study of Weiss and Kulikowski (1991) has tested a back-propagation model and concluded that the model is well suited for regression applications.

Classification Problem

Much marketing decision making involves the task of classifying an observation into one of several groups. Market researchers often face this type of problem in direct marketing and retail store site selection.

When companies adopt a strategy of direct marketing, they choose different methods such as catalogue-marketing and mail marketing to sell their products. In the case of mail marketing, marketers typically use their marketing database and make sales by mailing their products to the customers. They normally aim to secure immediate purchases from customers. The success of direct marketing is usually judged by the response rate. As a minimum response rate would cost the companies dearly, their objective is to maximize the response rate. It is possible to achieve a maximum response rate if the companies mail the product to customers who would be most able, willing and ready to buy (Kotler, 1991). Hence, one of the major decisions to be made in direct marketing is to decide the target/prospective customers.

The literature suggests different methods for identifying prospective customers (Stone, 1988). One method is to choose prospective customers by mail marketing to a sample of customers. In this method, companies mail their products to a sample chosen from their large customer database. Based on characteristics such as age, sex, income, occupation, education, social class and the geographic location of respondents in the testing, prospective customers can be chosen from a large database. Hence deciding on target/prospective customers is a problem of classifying customers in the database into prospective customers and non-prospective customers. Artificial neural networks could be used for identifying prospective customers. The conventional tool used for this purpose is discriminant analysis (Dillon and Mulani, 1989).

Discriminant Analysis

Discriminant analysis (DA) is a statistical technique that uses the available information on a set of variables for classification purposes. The analysis constructs a discriminant function for group variables on the set of independent variables. Numerous methods have been developed in constructing discriminant functions. They differ in two major aspects:

(1) assumptions regarding the group distribution;

(2) the functional form of the discriminant function.

The most commonly used form of discriminant function is of the linear type. For instance, in our insurance company example, if the managers like to predict the risk category (Z) of a new client based on age (X), sex (Y), income (W) and social class (V), then it calls for constructing the discriminant function of the form:

$$Z = a + b\,X + c\,Y + d\,W + e\,V$$

The above discriminant function can be constructed using Fisher's method (Hand, 1981), which maximizes the ratio of between-groups to within-groups variances. The discriminant function combines the information available in independent variables into a single valued estimate of Z that is called a discriminant score. As discriminant scores themselves are not generally equal to the values assigned to prospective and non-prospective customer groups, a classification rule is used to translate these scores into group membership predictions. A possible classification rule in this case could be the following: if the customer's discriminant score is less than or equal to some cut-off value C_1, then consider him/her as a prospective customer; otherwise consider him/her as a non-prospective customer. Determination of the appropriate cut-off value C_1 can be done in a number of ways (see Hand (1981)).

Difficulties in Discriminant Analysis

The above DA is optimal in minimizing misclassification, provided that each group follows a multivariate normal distribution and the covariance matrices of each group are identical. In practice, data pertaining to customer characteristics do violate the normality condition. Violation of the normality condition affects the predictive accuracy of the analysis (Huang and Lippman, 1987). Similarly, the violation of equal group variances affects the appropriate forms of the classification rule. It is important to realize that even a small improvement in choosing prospective customers would be beneficial to companies.

Neural Network Approach for Selecting Prospective Customers

Identification of prospective/target customers can be made using a neural network approach. A network can be constructed in which the number of nodes at the input layer is equal to the number of independent variables and the number of nodes at the output layer is equal to the number of groups. The ability of multi-layer networks to tackle any classification problem has been established in the literature.

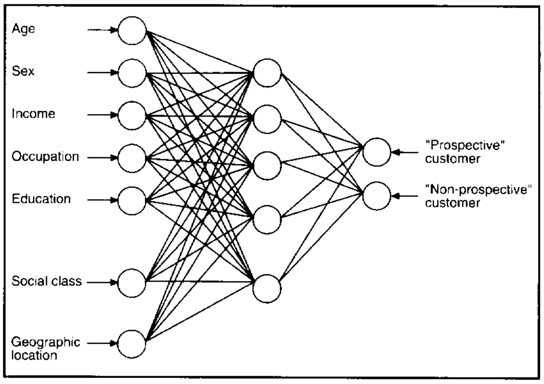
For example, in this case, a feed forward neural network with a single middle layer, consisting of seven input nodes, five middle-layer nodes and two output nodes (Figure 4) can be considered. For training the network, the back-propagation algorithm given can be used (Burke, 1991).

Data on the independent factors such as age, sex, income, occupation, education, social class and geographic location, and the group factors such as prospective and non-prospective customers, form a signal. Instead of using the integer values, say 1 and 2, at the output node representing prospective and non-prospective customers, all outputs are to be set to zero except for that corresponding to the class from which the input originates. That desired output is 1. The network can be trained by feeding the test signals over and over again. This would enable the network to adapt its weights, so that the network can classify the customers as "prospective" and "non-prospective". The processing that takes place in the network has already been described.

Neural Network Approach versus Discriminant Analysis for Identifying Prospective Customers

While handling a classification problem, neural networks make no assumption about the underlying statistical distribution in the data. They are more of a general method and perform better in classification tasks.

Figure 4. *7-5-2 Neural Network for Identifying Prospective Customers in Direct Marketing*



Like regression analysis, discriminant analysis also starts from scratch when a new case is added to the sample. This batch update may produce results provided that the distribution from which the new case is drawn is the same as the distribution of old cases. In situations where a new sample of several cases is drawn from a new distribution, these new cases are to be treated separately for classification purposes. Using old cases with the new cases may result in low predictive accuracy. On the other hand, neural networks adapt themselves as new examples become available and perform better classification. They do not ignore past information; instead, they reduce the importance of old cases as cases from new samples are fed into the network.

Various researchers (Huang and Lippman, 1987; Subramian *et al.*, 1993; Tam and Kiang, 1992; Yoon *et al.*, 1993) have compared the predictive power of neural networks with that of conventional discriminant analysis. All these experiments concluded that neural networks perform better than conventional discriminant analysis.

Grouping Problem

Another type of problem which market researchers/managers often face is the problem of grouping similar objects. For example, when companies are interested in test-marketing products, they face the problem of selecting "like" cities so that the results obtained are not attributable to a difference in market areas (Churchill, 1991). "Like" cities can be decided based on a number of characteristics such as population and median income. The other area in which market researchers face the grouping problem is target marketing.

Target marketing is a strategy that aims at grouping a major market into segments so as to target one or more of these segments or to develop products and marketing programmes tailored to each segment (Kotler, 1991). Using this strategy, companies can focus on the buyers whom they have the greatest chance of satisfying rather than scattering their marketing effort.

The first step in target marketing is market segmentation. Market segmentation is the process of dividing a market into distinct groups of buyers who might require separate products or marketing mixes. In simple words, it is the problem of grouping "similar" consumers.

There are many ways by which buyers/customers can be grouped. The basis for segmentation includes various characteristics of customers. In practice, customers' basic characteristics regarding demographics, socio-economic factors, geographic location and their product-related behavioural characteristics such as purchase behaviour, consumption behaviour and attitudes towards a product, are used for segmentation purposes (Dibb and Simkin, 1991). The number of segments being targeted will vary of course from market to market, for different products, and from company to company. Artificial neural networks could be used in determining segments. The conventional tool for solving such problems is cluster analysis.

Cluster Analysis

Cluster analysis is a statistical method for grouping. It is used to create groups, whereas discriminant analysis is used to assign cases to the existing groups. Groups may be achieved through a number of strategies, such as iterative and hierarchical approaches. Hierarchical methods are based on a "similarity measure" between objects. The most popular method in an iterative approach is the *k*-means approach. This method requires a number of groups, *k*, a priori. Seeds, the starting-point for those *k* groups, are to be determined either randomly or by some other means. In the first step, each city is assigned to one of the *k* groups. Then the average for each group is calculated based on the characteristics and the cities are reassigned on the basis of the mean to which they are closest. The above step is repeated until no objects are reclassified.

Difficulties in Cluster Analysis

Most cluster analysis methods are heuristic and are not supported by an extensive body of statistical reasoning (Aldenderfer and Blashfield, 1987). Most of the methods used for pattern recognition purposes have implicit assumptions about the underlying distribution of data. In addition to this, different clustering methods generate different solutions for the same data. On the other hand, in hierarchical clustering methods, researchers are always confronted with the problem of choosing a similarity measure. Hierarchical clustering methods are inappropriate when the sample size is large. Punj and Stewart (1983) have discussed the various issues that confront marketing researchers in using cluster analysis.

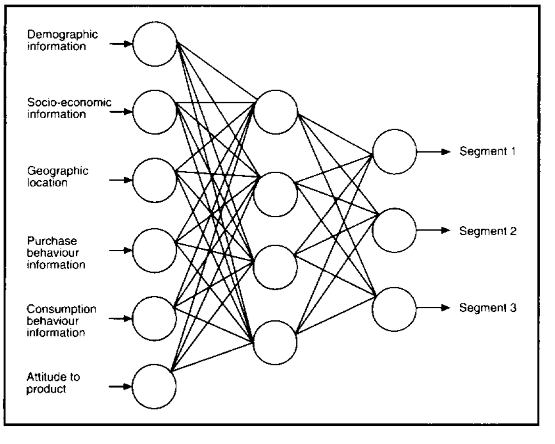
Neural Network Approach to Segmentation Problem

The division of a market into distinct groups of customers can be made using a neural network approach. A network can be constructed in which the number of nodes at the input layer is equal to the number of characteristics of customers considered, and the number of nodes at the output layer is equal to the number of segments required. A simple two-layer network (Figure 5) itself can perform the grouping task with unsupervised/competitive learning (Burke, 1991).

Data on customers' basic characteristics regarding demographics, socio-economic factors, geographic location and their product-related behavioural characteristics, such as purchase behaviour, consumption behaviour and attitudes to product, form a signal. Suppose the number of segments being targeted is three. The three segments are represented by three output nodes in the network. Of course, two-layer networks do not require knowledge about the precise number of segments being targeted. Signals can be fed one by one from the input node to each output node. In this process, the network computes the output at each output node. The output node with maximum value of output is said to "win" and the customer corresponding to the given input becomes the member of the "winning" node (segment).

Adaptive resonance theory (ART) and self-organizing feature maps (SOFM) (Lippman, 1987) extend competitive learning. They allow the network to assign an incoming signal to the segment having the nearest weight vector only if the distance between them falls within some predetermined limit; otherwise a new segment is formed.

Figure 5. *6-4-3 Neural Network for Segmentation*



Neural Network Approach versus Cluster Analysis Approach for Segmentation Problems

The main advantage of using neural networks for a segmentation problem is that they are robust. Another advantage is that, after segmentation, the same network can be used in future for classifying new buyers. This would often be the case when the company is seeking a better understanding of buyers' behaviours. Two-layer networks do not require knowledge of the precise number of segments being targeted. Instead, they adjust their parameters as the data are fed in and establish an arbitrary number of segments that represent the consumer data presented to the system. Parameters on the competitive learning model can be adjusted to tune the sensitivity of the system and to produce meaningful segments. Lippman (1987) has described the clustering aspect of different neural network models.

Conclusions

Owing to recognizing and learning abilities, neural networks can be applied to several marketing decision-making problems which were once reserved for multivariate statistical analysis. Though both neural networks and statistical techniques have common goals, there are many differences between these two techniques. The primary difference between neural networks and other statistical procedures is the method of processing data. In statistical techniques, processing is by batch and is sequential. Data are used only once. In neural networks, each datum in the sample is presented to the network repeatedly until the network identifies or learns the association of input to output. Moreover, this repeated processing takes place in a parallel and distributed way. Many processing elements share the job of working out the results.

Neural networks are capable of discovering the relationship, whereas regression/discriminant analysis requires knowledge of the nature of the underlying relationship.

Another feature of neural networks is that they are fault-tolerant. This means that the contribution made by any single processing element is not too important. Hence, even when there are missing elements in the sample data, the result/performance of the neural networks is not affected in a significant way. But, in statistical methods, problems with missing data are treated differently. For example, certain analysis requires correction terms to be added/used whenever there is a missing element. This correcting term is based on the number of missing elements.

Since the neural networks can estimate both quantitative variables (interval and ratio scale variables) and class variables (nominal variables), the same neural network and learning algorithm can be used for forecasting and classification purposes.

Also limitations of neural networks distinguish them from statistical techniques. For instance, formulae have been developed to determine the sample size for a given desired accuracy in statistical techniques. But there is no hard-and-fast rule in determining the sample size for training neural networks. A sample of larger size would lead to high accuracy, whereas a smaller sample would lead to low accuracy. When there is a severe constraint on sample size, it is suggested that one should increase the number of "epochs" (iterations) to improve efficiency. Likewise, other limitations of neural networks are the lack of explanation qualities and lack of a formal method to decide the network configuration for a given task.

In conclusion, we can say that neural network approaches differ from traditional statistical techniques in many ways, and the differences can be exploited by the application developer. They are powerful alternative tools and a complement to statistical techniques when data are multivariate with a high degree of interdependence between factors, when the data are noisy or incomplete, or when many hypotheses are to be pursued and high computational rates are required. With their unique features, both methods together can lead to a powerful decision-making tool. Studies and investigations are being made to enhance the applications of ANNs and to achieve the benefits of this new technology.

References and Further Reading

Aldenderfer, M.S. and Blashfield, R.K. (1987), *Cluster Analysis*, Sage Publications, London.

Burke, L. (1991), "Introduction to Artificial Neural Systems for Pattern Recognition", *Computers and Operations Research*, Vol. 18 No. 2, pp. 211-20.

Churchill, G.A. (1991), *Marketing Research: Methodological Foundations*, The Dryden Press, New York, NY.

De Groot, C. and Wurtz, D. (1991), "Analysis of Univariate Time Series with Connectionist Nets: A Case Study of Two Classical Examples", *Neuro Computing*, No. 3, pp. 177-92.

Dibb, S. and Simkin, L. (1991), "Targeting, Segments and Positioning", *International Journal of Retail & Distribution Management*, Vol. 19 No. 3, pp. 4-10.

Dillon, W.R. and Mulani, N. (1989), "LADI: A Latent Discriminant Model for Analyzing Marketing Research Data", *Journal of Marketing Research*, Vol. XXVI, pp. 15-29.

Eisenbeis, R.A. (1977), "Pitfalls in the Application of Discriminant Analysis", *Journal of Finance*, No. 32, pp. 875-900.

Hand, D.J. (1981), *Discrimination and Classification*, Wiley, New York, NY.

HNC (1990), "Neuro Software", HNC Inc., San Diego, CA.

Hruschka, H. (1993), "Determining Market Response Functions by Neural Network Modelling: A Comparison with

Econometric Techniques", *European Journal of Operations Research*, No. 66, pp. 27-35.

Huang, W.Y. and Lippman R.P. (1987), *Comparisons between Neural Net and Conventional Classifiers*, IEEE 1st International Conference on Neural Networks IV, pp. 485-93.

Jurik, M. (1993), "Consumers' Guide to Neural Network Software", *Futures: The Magazine of Commodities and Options*, pp. 36-41.

Kohonen, T. (1988), *Self-organisation and Associative Memory*, Springer, New York, NY.

Kotler, P. (1991), *Marketing Management*, Prentice-Hall, Englewood Cliffs, NJ.

Lippman, R:P. (1987), "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine*, April, pp. 4-22.

Nelson, M.M. and Illingworth, W.T. (1991), *A Practical Guide to Neural Nets*, Addison-Wesley, New York, NY.

Proctor, R.A. (1992), "Marketing Decision Support Systems: A Role of Neural Networking", *Marketing Intelligence & Planning*, Vol. 10 No. 1, pp. 21-6.

Punj, G. and Stewart, D.W. (1983), "Cluster Analysis in Marketing Research: Review and Suggestions for Application", *Journal of Marketing Research*, Vol. XX, pp. 134-48.

Stone, B. (1988), *Successful Direct Marketing Methods*, 4th ed., Lincolnwood, IL.

Subramanian, V., Hung, M.S. and Hu, M.Y. (1993), "An Experimental Evaluation of Neural Networks for Classification", *Computers and Operations Research*, Vol. 20 No. 7, pp. 769-82.

Tam, K.Y. and Kiang, M.Y. (1992), "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions", *Management Science*, Vol. 38 No. 7, pp. 926-47.

Thall, N. (1992), "Neural Forecasts: A Retail Sales Booster", *Discount Merchandiser*, Vol. 32 No. 10, pp. 41-2.

Weiss, S.M. and Kulikowski, C.A. (1991), *Computer Systems That Learn*, Morgan Kaufmann, CA.

White, H. (1989), "Neural Network Learning and Statistics", *AI Expert*, Vol. 4 No. 12, pp. 48-82.

Yoon, Y., Swales, G. and Margavio, T.M. (1993), "A Comparison of Discriminant Analysis versus Artificial Neural Networks", *Journal of Operational Research Society*, Vol. 44 No. 1, pp. 51-60.

V. Venugopal is a Visiting Faculty Member and W. Baets is an Assistant Professor at The Netherlands Business School, Nijenrode University, Breukelen, The Netherlands.