Forecasting and Analysis of Marketing Data Using Neural Networks

Jingtao Yao Nicholas Teng Hean-Lee Poh Chew Lim Tan
Department of Information Systems and Computer Science
National University of Singapore
Singapore 119260
email : yaojt@comp.nus.edu.sg
Tel: (65)874 6729
Fax: (65)779 4580

First Draft: May 1996, 1^{st} Revision: June 1997, 2^{nd} Revision: October 1997
Journal of Information Science and Engineering 14, No. 4, pp523-545(1998)

Abstract

This study aims to incorporate Artificial Neural Networks into a Marketing Decision Support System (MDSS), specifically, by discovering important variables that influence sales performance of colour television (CTV) sets in the Singapore market using neural networks. Three kinds of variables, expert knowledge, marketing information and environmental data, are examined. The information about the effects of each of these variables has been studied and made available for decision making. However, their combined effect is unknown. This study attempts to explore the combined effect for the benefit of our collaborator, a multinational corporation (MNC) in the consumer electronics industry in Singapore. Putting three kinds of variables together as input variables results in a neural network model. Neural network training is conducted using historical data on CTV sales in Singapore collected over the past one and a half years. Sensitivity analysis is then performed to reduce input variables of neural networks. This is done by analyzing the weights of the input node connections in the trained neural networks using two different methods. The weaker variables can be excluded and this results in a simpler model. Further, an R-Square value of almost 1 is obtained through the inclusion of an Unknown variable when the network model consisting only of the most influential variables is trained and tested. Knowing the most influential variables, which in this case include Average Price, Screen Size, Stereo Systems, Flat-Square screen type and Seasonal Factors, marketing managers can improve sales performance by paying more attention to them.

Keywords: *Artificial Neural Networks, Marketing Decision Support Systems, Sales forecasting, Marketing mix, Variable reduction.*

1 Introduction

This study aims to incorporate Artificial Neural Networks in a Marketing Decision Support System (MDSS). Specifically, its aim is to discover important variables that influence sales performance of colour television (CTV) sets in the Singapore market using neural networks. Surviving in today's business environment calls for faster and better decisions. Marketing decision makers are increasingly drawing on computer-based Decision Support Systems (DSS) to help marshal the vital information that they need to make informed choices.

Little [1] stressed the importance of supporting marketing managers with good information. In order to compete, information is an essential tool. With the provision of good information, the MDSS seeks to help make better informed, more timely and more effective marketing decisions. The benefits of effective and timely decisions are not always tangible, but these are the main reasons for the existence of an MDSS. Sales are influenced by many factors, such as price, product, promotion and place (also well known as the 4-Ps); hence, they are very difficult to forecast.

Neural Network technology can be applied in many areas especially when the problem domain involves classification, recognition and prediction. A recent survey of research [2] conducted by Wong etc. indicated that at least 127 business application journal papers had been published up to September 1994. Looking at the marketing data that a marketing manager has to deal with, it is not difficult to see that he has to sift out useful information from raw data. Considering the multi-faceted demands of the marketing domain and the myriad sets of available marketing data, it can be very difficult to draw anything relevant from the available data because of the unrelatedness and non-linearity of marketing data. With the capabilities of neural network, hidden trends and relations among data which are previously unseen can be deduced, in other words, obtaining information from information [3]. In the light of such capabilities of neural networks, the power of MDSS will be greatly enhanced if neural networks can predict sales performance from previous sales and marketing data.

This article is organized as follows. The second section presents neural networks and describe how they can be incorporated into an MDSS. In the third section, background information on data collection is provided. The fourth section presents the variable selection approach and forecast model. The fifth section details the sensitivity analysis used to reduce the number of input variables. Research findings and practical implications are included in the sixth section. Some concluding remarks are offered in the final section.

2 Neural Networks in Marketing

A neural network is a collection of interconnected simple processing elements. Every connection in a neural network has a weight attached to it. The backpropagation algorithm [4] has emerged as one of the most widely used learning procedures for multi-layer networks. It wss selected for use in this study because of its suitability for applications which involve predictions. The typical backpropagation neural network usually has an input layer, some hidden layers and an output layer. The nodes in the network are connected in a feedforward manner, from the input layer to the output layer. The weights of connections are given initial values. The error between the predicted output value and the actual value is backpropagated through the network to update weights. This is a supervised learning procedure that attempts to minimize the error between the desired and the predicted outputs. The relationship can be obtained through a group of mappings with constant time interval.

The key feature of this kind of networks is that it is able to map an extracted pattern from the input stimuli to the output; hence, it is good for classification. This is because the nodes in the hidden layers are able to learn to respond to uniqueness in the input layer. This uniqueness is actually the correlation activities of the input layer nodes [5]. These nodes can affect each other; e.g., the first node can affect or be correlated to the last node. Such relationships between the input nodes provide a basis for more abstract representation of the input information in the next higher layer.

As the network are trained under supervision with different examples provided, it acquires the ability to generalize. This ability is trained into the hidden layer nodes during the training so that when previously unseen patterns are presented, the network is probably able to recognize and classify that pattern accordingly. In order for a feedforward network to perform successfully, its ability to recognize features amongst the input nodes must be ensured. This is where the hidden nodes must be trained to recognize the general features of the input pattern. Such features must be sufficiently general so that the network is not led into overfitting.

Considering the multi-faceted demands of the marketing domain and the myriad sets of available marketing data, it can be very difficult to draw anything relevant from the available data because of the unrelatedness and non-linearity of marketing data. A study was carried out by Proctor [6] on the role of neural networks in marketing. The strength of a neural network is its ability to obtain the relationships of non linearly dependent variables. This was further emphasized by Dutta *et al.* [7], who incorporated neural networks into a Decision Support System (DSS). In a separate work, Venugopal and Baets [8] also proposed using neural networks in retail sales forecasting, direct marketing and target marketing.

Compared with other traditional statistical methods, neural networks require only minimum knowledge to the problem's structure [9]. No prior knowledge of the statistical distribution of the data is required because the network develops an internal relationship between the variables [8]. All these are combined to make neural networks particularly suitable to complex classification problems in which the mapping rationale is either fuzzy, inconsistent, or completely unknown. Most business and marketing applications can be considered to be classified under fuzzy classifications [9], and this is the area that neural networks can be of service. When used for fuzzy classification, neural networks can be said to be able to see through "noise".

As stated by T. Hill [10], there are three inherent limitations for regression MDSS models. First, they are based on inear combinations of decision variables only; second, human understanding is necessary specify the model or data transformations; third, there are no learning facilities although new situations arise and the results need to be estimated frequently. Neural Networks do not suffer from the limitations of regression models and have been proven to be able to learn functional relationships from input variables to predict results [3] [7] [10] [11].

In our experiments, one hidden layer backpropagation neural networks were used. The number of nodes in a hidden layer ranged from one to two-thirds of the number in input layers. The activation function used was a sigmoid

function.

3 Background Information

In Singapore, consumer electronics is a very competitive market. Big established brands command a certain loyalty from consumers. Consumers in the market tend to go for brands that are familiar, even though comparisons show no difference in technology. Knowing the brand but not always the quality, consumers will usually replace old electrical appliances with newer models of the same brand or models from other the familiar brands.

With the help of a multi-national corporation in Singapore, this study was conducted. This company deals with consumer electronics and has been trying to penetrate the local market for some time. It has not been particularly successful, and their main strategy up to now has been to gain as much market share as possible. The nature of the market makes operations very difficult. The product life cycle is rather short, about one and a half to two years for a typical model, and the prices are very inelastic. In order for a marketing strategy to be successful, careful planning is required. This can be done by having a good combination of the 4-Ps of the marketing mix. An effective mix can be discovered.

This study implemented neural network technologies to discover an effective marketing mix of the 4-Ps of marketing. This was done by extracting variables that have more influence on Sales Revenue and excluding the less influential ones. Using neural networks, another goal of this study was be to provide information on the relative influence of each variable in the marketing mix. The marketing variables were used as independent variables, and the dependent variable was simply the *Sales Revenue* of colour television sets in Singapore. In the process, it was hoped that the neural network component of the company's MDSS could also be designed.

4 Variables Selection Methodology and Models

This study looked into the combined effects of data on Sales Revenue. Sales Revenue was assumed throughout this study to be the marketing objective of the company's marketing plan. The data on CTV sales in Singapore covers 4 bi-monthly periods from March 1994 to April 1995 (See Table 1). They were obtained from the company and some publications (e.g., World Economic Fact book). The data were separated in three groups: expert knowledge, marketing and environment.

Following Kotler's marketing model [12], 20 independent variables were selected (See Table 2) in this study. Since the percentage of urban population and average household size do not vary significantly in the Singaporean context, they were not used for predictions in Singapore. Finally, 18 input variables and one output variable were selected for the initial data model.

Bi-Monthly Periods	Data Rows Contributed
March - April 1994	1 to 66
August - September 1994	67 to 124
December 1994 - January 1995	125 to 187
April - May 1995	188 to 242

Table 1: Four Bi-monthly periods from which data were collected on the CTV market in Singapore.

Sales Promotion, Seasonal Factors and *Brand Preference* are special variables that model the experience or "expert knowledge" of marketing managers. These variables affect sales in a significant way. *Sales Promotion* is used to indicate the relative enthusiasm of show-room promoters in promoting television sets. Such enthusiasm is directly triggered by the amount of the commissions they receive for selling a particular television set. Therefore, brands that give commissions are preferred by sales promoters more than brands that do not. Sales promoters will generally introduce brands to buyers that give more commissions.

Seasonal Factors are reasons for the fluctuation of the sales of television sets in a year. It is noted that sales tend to peak especially during seasonal holidays and school vacations. These seasons are November to January(school holiday, Chinese New Year), late February to March(Chinese New Year, Hari Raya Puasa and short school holiday) and June(school holiday). A value of 1 is used to indicate that a datum is obtained in season, while 0 indicates that a datum is obtained out of season.

Brand Preference is used to indicate the preferred brands of television sets that Singaporean buy. This is a factor that any new player into the market has to contend with. The preferred brands are generally brands that have been established in Singapore for a period of time. They have proven themselves to be brands that are trusted by the people. Such brands now command premium prices. Loyalty to such brands is strong, so to persuade buyers to buy new brands, it is a necessary to break this old "trust".

There are altogether 242 sets of data available for the analysis. Since neural networks only accept real numbers, variables with alphabetic values have to be converted into numerical values. For example, the variable *System* has three different values : MONO, NICAM and German NICAM. The numbers 1, 2 and 3 are used to represent these three system types, respectively.

Variable Names	Variable Group	Data Type	Mean	Standard Deviations	4-Ps
Average Price	Marketing	Continuous	S$1,063.97	S$800.14	Price
System	Marketing	Discrete	2.2851	0.9364	Product
Screen Size	Marketing	Continous	22.0496″	6.0761″	Product
Flat-Square	Marketing	Discrete	2.0785	0.7010	Product
Stereo	Marketing	Discrete	1.88430	0.9484	Product
Text	Marketing	Binary	0.6160	0.4870	Product
Sales Promotion	Expert Knowledge	Binary	0.9380	0.2416	Promotion
Press	Marketing	Continuous	S$301,030	S$147070	Promotion
Magazine	Marketing	Continuous	S$11,690	S$17840	Promotion
TV	Marketing	Continuous	S$377,320	S$292600	Promotion
Bus/Taxi	Marketing	Continuous	S$58,302	S$62770	Promotion
Inflation	Environmental	Continuous	3.2%	0.2004%	Place
GDP	Environmental	Continuous	$63,195.1 Billions	$2729.034 Billions	Place
Population Size	Environmental	Continuous	2.9493 Millions	0.0301 Millions	Place
Product Penetration	Environmental	Continuous	97.9636%	0.0551%	Place
Seasonal Factor	Expert Knowledge	Binary	0.26	0.44	Place
Brand Preference	Expert Knowledge	Discrete	2.6160	0.615	Place
Market Demand	Marketing	Continuous	S$215,847,000	S$23,917,000	Place
Urban Population Size	Environmental	Continuous	100%	0	Place
Average Household Size	Environmental	Continuous	3.5 persons	0	Place

Table 2: Summary of variable data collected for analysis from four bi-monthly periods from March 94 to May 95, their sources and relationship to the 4-Ps of marketing. Variable Group refers to the three different groups of variables the marketing manager uses to make decisions.(The statistic may not be meaningful for binary or discrete data.)

The neural network model has the task in the marketing decision support system of forecasting the quantity or sales revenue as marketing managers wish. Its inputs are determined by the marketing mix that contributes to the sales performance of a product in terms of quantity sold or sales revenue.

Backpropagation neural networks were used in this study to capture the relationship between these marketing variables and sales revenue. For most cases, one hidden layer networks were used following the authors' experiences. This can be considered as a rule of thumb. The number of nodes used in hidden layers is around half of the number of input variables of a neural network. The maximum number of training iterations was set to 10,000 while the learning rate and momentum rate ranged from 0.001 to -0.3 and 0.5 to 0.8, respectively.

5 Sensitivity Analysis Through Analysis of Weight

Of the 18 variables selected, preliminary tests were run in an attempt to discover which of these variables influenced the output variable. As a rule of thumb to determine whether a variable is relevant, a network was run numerous times, each time omitting one variable. If the results before and after omitting a variable were the same or even better, it could be inferred that this variable probably did not contribute much to producing the outcome. Such variables included *Penetration, System, Stereo* and *Magazine* among others. On the contrary, if the results of the network deteriorated significantly after the variable had been left out, then the variable probably had a big influence on the outcome. Such variables included *Average Price, TV, Bus/Taxi, Sales Promotion* and *Screen Size.* The findings with respect to *stereo* were somewhat surprising as it is known to have been a much sought after feature in the society of Singapore in recent years. To investigate this further, weight analysis, which is the subject of the next section, was conducted. It could throw light on the relative influence of each variable on the output variable.

The R-Square [13] was used as the basis for comparing the test results of all the test runs throughout this study. It is a measure of how well the outcome (i.e., the dependent variable) of the network is described by the independent variables (i.e., input variables). The R-Square value is obtained by the equation:

$$RSquare = 1 - NMSE, \tag{1}$$

where $NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} = \frac{1}{\delta^2}\frac{1}{N}\sum_k (x_k - \hat{x}_k)^2$, which is popularly used for error measurement [14] [15],

and, x_k: the actual target value;
$\hat{x}_k$: the output produced by the neural network;
$\bar{x}_k$: the mean of x_k;
δ^2: the variance of the data set.

The RSquare or NMSE can be used to measure the average fitness of the forecast values and target real values. In addition, the correlation coefficients between the forecast time series and target real time series are used to judge the trend fitness of the two series. If the fitness between each pair of points and their trend is good, we then can say that the model we have used is one of the best among the models we have experimented.

5.1 Three Sensitivity Methods

A good and understandable data model can be provided by reducing the number of variables through weight sensitivity analysis [16]. The analysis of weights can be accomplished using three methods proposed in this study. The idea of continuity of variables aids us to choose the methods for analysis. This is because a variable might not be meaningful when it is analyzed based on a certain method used for analysis.

The first method is the *Equation Method.* For a feedforward neural network with one hidden layer and one output node, the influence of each input variable on the output can be calculated by the equation:

$$\sum_k O(1-O)w_{k1}^2 v_k^2 (1-v_k^2) w_{ik}^1 \forall i, \tag{2}$$

where O: the value of the output node;
w^2_{k1}: the outgoing weight of the kth node in the hidden (2nd) layer;
v^2_k: the output value of the kth node in the hidden (2nd) layer; [1]
w^1_{ik}: the connection weights between the ith node of the input (1st)layer and the kth node in the hidden layer.

For the *Equation Method*, there will be n readings for n input variables for each input row into the network. If there are r input rows, there will be r readings for each of the n input variables. All the r readings for each input variable are subsequently plotted to obtain its mean influence, I_n, on the output. These 'I_n's indicate the relative influence each input variable can have on the output variable; the greater the value, the higher the influence.

The second method is the *Weight Magnitude Analysis Method.* The connecting weights between the input and the hidden nodes are observed. The rationale for this method is that variables with higher connecting weights between the input and output nodes will have greater influence on the output node results. For each input node, the sum of its output weight magnitudes to each of the hidden layer nodes is the relative influence of that node on the output. To find the sum of the weight magnitudes from each input node, the weight magnitudes of each of the input nodes are first divided by the largest connecting weight magnitude between the input and the hidden layer. This is called normalization. The normalization process is a necessary step whereby the weights are adjusted in terms of the largest weight magnitude. The weight magnitudes from each input node to the nodes in the hidden layer are subsequently summed and ranked in a descending manner. The rank is an indication of the influence that an input node has on the output node relative to the rest. The rank formula is as follows:

$$I_i = \sum_k \frac{w^1_{ik}}{max_{All\ i,k}(w^1_{ik})} \tag{3}$$

where the notation is the same as in Equation 2.

The third method is the *Variable Perturbation Method.* It is used to adjust the input values of one variable while keeping all the other variables untouched. These changes can take the form of $I_n \Rightarrow I_n + \delta$ or $I_n \Rightarrow I_n * \delta$, where I is the input variable to be adjusted and δ is the change introduced into I. The corresponding responses of the output against each change in the input variable are noted. The node whose changes affects the output most is the one that is most influential relative to the rest. Under the *Variable Perturbation Method*, the binary variable *Seasonal Factors* will have values of 1.1 or 0.9 when it is 10% perturbed, but the values are not meaningful for this variable. *System*, a discrete variable on the other hand, has values of $1, 2$ or 3. It will not be meaningful if it has values of $1.1, 2.2$ or 3.3 under a 10% perturbation. Thus, the need to analyze the variables according to their continuity is there.

5.2 Deduction of Continuous, Binary and Discrete Variables

The variable reduction approach was used first to separate all the variables into their respective categories in accordance with the continuity of the values: binary, continuous and discrete. The influence of each variable on the outcome was then ranked within the respective categories. Thereafter, the best few ranking variables within each category were selected and combined. Runs for the combined variables were done, and the results (especially R-Square) were noted and used as a benchmark.

[1] The notation w^a_{bc} here can be read as the weight from the bth node in the ath layer to the cth node in the next layer.

This variable combination for the benchmark was the best one until another combination produced better results, and the benchmark then took on these new results. This cycle was repeated until the benchmark could not be further improved.

The variable combinations change in numbers or the selection of different variables while the number remained constant. This was done by using one of the weight analysis methods mentioned above. To reduce the numbers, the best few variables of each run were selected for further reduction while the rest were discarded. This went on until discarding even the least influential variable might not improve the benchmark and might produce inferior results. Variable selections could be changed by excluding these remaining variables one at a time in order to find a combination that could improve the benchmark. If no such new combinations were possible, the last selection that produced the benchmark was most representative of the data model. The final data model still represented the problem meaningfully, even after the weaker variables were removed.

The continuous variables were examined, and those with weaker influence on the outcome were discarded. The ranking of the variables in terms of their influence on the outcome is presented in Table 3. Referring to the preliminary results, it is found that *Screen Size* is consistently regarded as the most significant variable. On the other hand, *Bus/Taxi* and TV are again considered as one of the least significant variables. *Inflation*, however, is only consistent within the method but varies slightly when compared with the other method. Nevertheless, even though the two methods did not produce identical rankings, the relative influence of the variables can be perceived by observing that the rankings are very similar between the two methods.

We next reduced the number of continuous variables to the five most influential variables. Based on the rough standing in Table 4 and 5, two sets of seven variables were selected from the two tables, respectively. Neural network tests were conducted on them to find their effects on the outcome. The new ranking is shown in Table 6. The two lowest ranking variables(*Magazine* and *GDP*) from each group were discarded. We next combined the two groups of variables to conduct another round of experiments and picked the seven top variables from the combined list. The seven chosen variables are *Average Price, Press, Screen Size, Inflation, Population Size, Penetration* and *Market Demand.* As the two variables, *Screen Size* and *Average Price* were ranked "1" within their groups as shown in Table 6, we considered them to be the two most influential variables. Different combinations of five variables were then formed from the seven chosen variables, with each combination always containing the two most influential variables. We subjected each combination together with the discrete and binary variables to neural network testing. The results of such runs of 12-variable (5 continues + 3 binary + 4 discrete) networks are shown in Table 7, where combination 2 produced the best results and was selected. In conclusion, the five continuous variables, *Screen Size, Average Price, Press, Market Demand* and *Population Size*, emerged as the five most influential continuous variables.

After the continuous variables were tested, the next test to be performed was to reduce the binary variables. They were *Sales Promotion, Text* and *Seasonal Factors.* These three variables were tested by excluding them one at a time from the other 15 variables. The steps followed closely the variable perturbation method. Since they were binary, the values of the variables to be tested were inverted and tested while the other variables were kept untouched. Results indicated that these three variables affected the outcome and needed to be included until they were combined with the discrete and continuous variables.

Following the binary variables, the influence of discrete variables was tested. There were only

	Weight Magnitude Method				Variable Perturbation Method on 18 variable networks			
Variable Names	18 var network	18 var. with Unknown network	16 var § network	16 var § with Unknown network	-10% †	-5% †	+5% †	+10% †
Average Price ‡	4	7	1	3	8	9	10	8
Press ‡	2	3	4	7	7	6	7	2
Magazine	8	11	-	-	10	2	3	3
TV	10	12	6	11	8	11	8	11
Bus/Taxi	11	10	9	9	9	10	9	9
Screen size ‡	1	2	2	2	1	1	1	1
Inflation	3	4	3	4	5	4	6	6
GDP	5	5	8	5	4	7	4	4
Population Size	6	8	7	8	3	3	2	5
Penetration	7	6	10	6	6	11	11	7
Market Demand ‡	9	9	5	10	2	5	5	10
Unknown	-	1	-	1	-	-	-	-

Table 3: Reduction of continuous variables and a comparison between the Weight Magnitude and the Variable Perturbation methods of weight analysis. The numbers represent the relative influence on the outcome. These numbers are the rankings in descending order obtained by the corresponding methods under the corresponding conditions or variable mix. §: Two variables, magazine and stereo, are excluded because of the results in the previous section, where they were shown to not have any influence on the outcome. †: The four tests here were done inclusive of Screen Size. For this variable, the variation $-2''$ corresponds to -10% perturbation, $-1''$ to -5%, $+1''$ to $+5\%$ and $+2''$ to $+10\%$, respectively. ‡ : The most influential continuous variables

four discreet variables to be reduced: *Brand Preference, System, Flat-Square* and *Stereo*. These four variables with the most influential five continuous variables and all three binary variables were put together to produce a network of twelve input variables and an output variable. The reason for this combination was to reduce the training effort since there were not many discrete variables. See Table 8 for the twelve variables and the training results of this consequent network. The best five variables were found after several deduction steps. The results in Table 9 show that no more deduction could be conducted.

5.3 Final Model and Influence with Unknown Factor

Referring to Table 3, it is noteworthy that in the cases where the *Unknown* variable was included, it was ranked the highest. The *Unknown* variable was created by a random generator. Both cases indicated that there were still some unknown factors affecting the outcome that were not captured by the known variables. The effect of *Unknown* had on the results of the final variable mix is demonstrated in Table 10. The *Unknown* was necessary as there could be too many interactions amongst the variables that cumulatively weakened the influence of individual inputs on the outcome. The *Unknown* enabled the network to capture and absorb these miscellaneous influences. The addition of variables meant that there have been even more interactions among the variables, although the intention might be to better explain the unknown factors described by the *Unknown*. Conversely, the subtraction of

Variable Name	18 var. Network	18 var. with Unknown Network	16 var. Network	16 var. with Unknown Network	Average Ranking	Rough Standing
Average Price §	4	7	1	3	3.75	3
Press §	2	3	4	7	4	4
TV	10	12	6	11	9.75	9
Bus/Taxi	11	10	9	9	9.75	9
Screen Size §	1	2	2	2	1.75	1
Inflation §	3	4	3	4	3.5	2
GDP §	5	5	8	5	5.75	5
Population Size§	6	8	7	8	7.25	6
Penetration §	7	6	10	6	7.25	6
Market Demand	9	9	5	10	8.25	8
Unknown	-	1	-	1	-	

Table 4: Rough standings of each variable and the best ranking seven variables from the Weight Magnitude Method. §: Chosen Variables

Variable Name	-10%	-5%	+5%	+10%	Average Ranking	Rough Standing
Average Price	8	9	10	6	8.25	8
Press ‡	7	6	7	2	5.5	6
Magazine ‡	10	2	3	3	4.5	3
TV	8	11	8	11	9.5	11
Bus/Taxi	9	10	9	9	9.25	10
Screen Size ‡	1	1	1	1	1	1
Inflation ‡	5	4	6	6	5.25	5
GDP ‡	4	7	4	4	4.75	4
Population Size‡	3	3	2	5	3.25	2
Penetration	6	11	11	7	8.75	9
Market Demand ‡	2	5	5	10	5.5	6

Table 5: Rough standings of each variable and the highest ranking seven variables from the Variable Perturbation Method. ‡: Chosen Variables

variables from the current selection might have removed some of the detrimental mutual interactions amongst the variables and in doing so, might produce better results. A reduced number of variables would make the model more understandable, and time could be saved through efficient training of the neural network.

After the best five variables were found, a *Unknown* was included in the final data model to examine the effects of such inclusion. The *Unknown* is ranked as most influential on the outcome in Table 10. This implies that many extra factors were not captured by the five variables. All these factors were trapped by the *Unknown*, leading to the good results reflected in Table 10 and Figure 1. Based on the ranking in this table, the relative influence of each of the variables could be put in perspective when the *Unknown* was present (See Table 1). It can be concluded that with the given data and under the method used for analysis, the groups of variables shown in Table 10 produced the best predictions of the *Total Sales Revenue* of the sales of colour television sets in Singapore. The output variable, *Total Sales Revenue*, was most sensitive to changes in these five variables. Conversely, changes in these

Weight Magnitude Method		Variable Perturbation Method	
Variable Name	Position Within Group	Variable Name	Position Within Group
Average Price	1	Press	3
Press	4	Magazine	7
Screen Size	2	Screen Size	1
Inflation	3	Inflation	2
GDP	7	GDP	4
Population Size	5	Population Size	5
Penetration	6	Market Demand	6

Table 6: The seven chosen variables and their ranks under both ranking methods. The ranks were obtained by running network tests on only the 7 variables of the respective groups. GDP and Magazine were ranked 7th by the corresponding methods and were discarded.

Combinations	Continuous Variables Chosen	R-Square of Testings
1	Screen Size, Ave Price, Infln, Press, MktDmd	0.23905
2	**Screen Size, Ave Price,Press, MktDmd, Population**	**0.71050**
3	Screen Size, Ave Price,MktDmd, Population, Penetration	< 0.001
4	Screen Size, Ave Price, Population, Penetration, Infln	< 0.001
5	Screen Size, Ave Price,Penetration, Infln, Press	0.61963
6	Screen Size, Ave Price, Press, MktDmd, Penetration	0.23525
7	Screen Size, Ave Price, Population, Infln, Press	< 0.001
8	Screen Size, Ave Price, Infln, MktDmd, Penetration	0.67007
9	Screen Size, Ave Price, Infln, MktDmd, Population	0.47251
10	Screen Size, Ave Price, Penetration, Press, Population	0.58112

Table 7: Results of tests to select the best 5 continuous variables. The variables chosen are joined with binary and discrete variables.

variables were most likely to influence *Total Sales Revenue*. It is not possible, however, to pinpoint the magnitude of influence one variable had on the outcome over the other variables. Judging from the *Unknown*'s weight magnitude, it absorbed a fair amount of unaccounted and unknown factors.

Based on the methods used in this paper, the aim was to separate the more influential variables from those that were less influential. Another way to look at the variable reduction process is to discover the weaker variables, variables with no influence and variables whose information is duplicated in other variables. Finding variables whose information is duplicated requires that the network outcome be almost identical when this variable and its duplicate are interchanged. It should be kept in mind that there is a possibility that more than one double of a variable could be present. Once it is discovered that the network produces the results when either a variable or its duplicate(s) is(are) absent, only one of them is required and the rest can be removed.

Figure 2 depicts the final neural network model obtained by this research. Sixty rows of input data were collected from November 1995 on the Singapore colour television market. They were fed into the trained network with *Unknown*, and the R-square and coefficient of correlation were 0.99577 and 0.99834, respectively. The prediction is shown in Figure 3.

Variable Name	Ranking	Continuity
Average Price	1	continuous
Press	11	continuous
Sales Promotion	5	binary
System	9	discrete
Screen Size	4	discrete §
Flat-Square	3	discrete
Stereo	2	discrete
Text	12	binary
Population Size	10	continuous
Seasonal Factor	6	binary
Brand Preference	8	discrete
Market Demand	7	continuous
R-square of testing	0.7105	
Coefficient of Correlation	0.89238	
R-square of testing (16 var)	0.65988	
Coefficient of Correlation (16 var)	0.86451	

Table 8: Training results of the 4 discrete variables combined with the 5 best continuous variables and the 3 binary variables. §: Screen Size was grouped together with the continuous variables. This model wea taken as the initial benchmark.

Variable Removed	Correlation Coefficient	R-square of testing
Seasonal Factor	0.62584	0.35716
Stereo	0.59189	<0.001
Flat Square	0.61499	0.40214
Screen Size	0.61499	0.39621
Average Price	0.67467	0.48167

Table 9: Results of 4-variable networks with one variable excluded at a time from the 5-variable benchmark combination.

6 Findings and Practical Implications

The surviving variables obtained were: *Average Price, Screen Size, Flat-Square, Stereo*, and *Seasonal Factors*. These variables were most influential on the *Sales Revenue*. Conversely, *Sales Revenue* was most sensitive to these factors. Although there is no systematic way of deducing the proportion of influence these variables have, they are factors to be considered when deciding on a marketing strategy.

The inclusion of *Average Price* in these final variables indicates that the consumers in Singapore are price sensitive. Therefore, when setting prices, a company should be careful not to set price too high. Pegging the prices to those of established brands would be an suggestion. However, doing so might make the television sets of the company seem overpriced since the established brands sell at premium prices. To be able to join the league of established brands, the price might be pegged at a percentage lower at the same time, advertising could stress that the television sets surpass the competitors' sets in prestige and recognition in other parts of the world, say, Europe or America. In order to penetrate the Singapore market, prices should not be set too High; if they are properly, some consumers might just switch brands.

There is a certain preference for *Screen Size* in Singapore. Most families are moving to a screen size

Variable Name	Ranking	Continuity
Average Price	3	continuous
Screen Size	2	discrete
Flat-Square	6	discrete
Stereo	5	discrete
Seasonal Factor	4	binary
Unknown	1	-
R-square of testings	0.99781	
Coefficient of Correlation	0.99889	

Table 10: The Effect of *Unknown* on the Benchmark 6-variable Network.

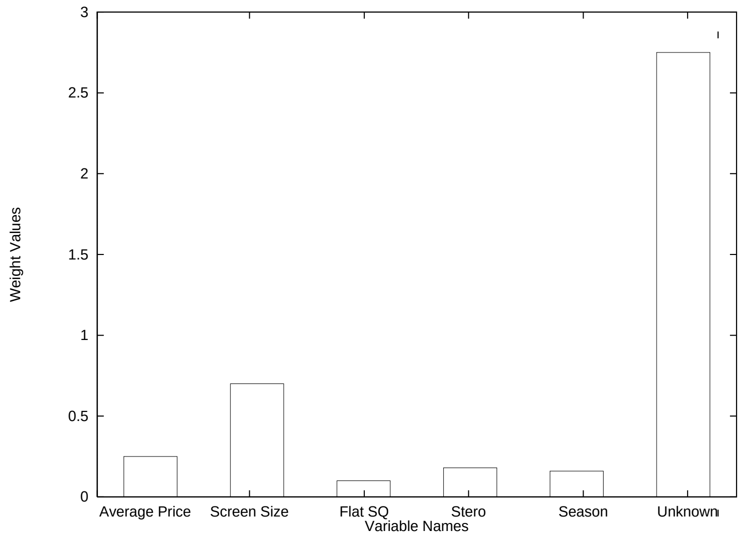

Figure 1: The effect of *Unknown* on the final 5-variable combination. Relative influence of input node weight magnitudes on the outcome.

that is currently 29 inches. As this factor plays a part in influencing sales, the management might need to be careful to ensure that their sets in this screen size are appealing enough to consumers. If the management is conservative, it will want to find out where this trend is heading and be prepared to introduce and promote sets of the this screen size before its competitors do. An adventurous management team might want to set the trend by promoting a new screen size and advertise it as the next "in thing". Doing so would avoid confronting the established Brands, which are strongly entrenched in such screen sizes.

The variable *Flat-Square* reflects the trend of consumers to be attracted by looks and technology. The flat and square screen is at present, the "sought after feature" in a television set if the price is right. Though consumers might not be aware that flat-square tubes cost more and are technologically more demanding, they do not mind paying more if the technology delivers some intrinsic value, such as aesthetic feelings, and quality pictures. If the technological aspects are investigated further, this might reveal some hidden desires of general television buyers to acquire sets that are more technologically advanced. This desire could also be a wish to own and enjoy finer things in life. Good picture and sound quality complement each other. It is not surprising that the *Stereo* feature of a set should, together with *Flat-Square*, be considered as among the most influential factors. Sets with *Stereo* and *Flat-Square* features sell for higher prices and consumers might not consider buying a set if one feature

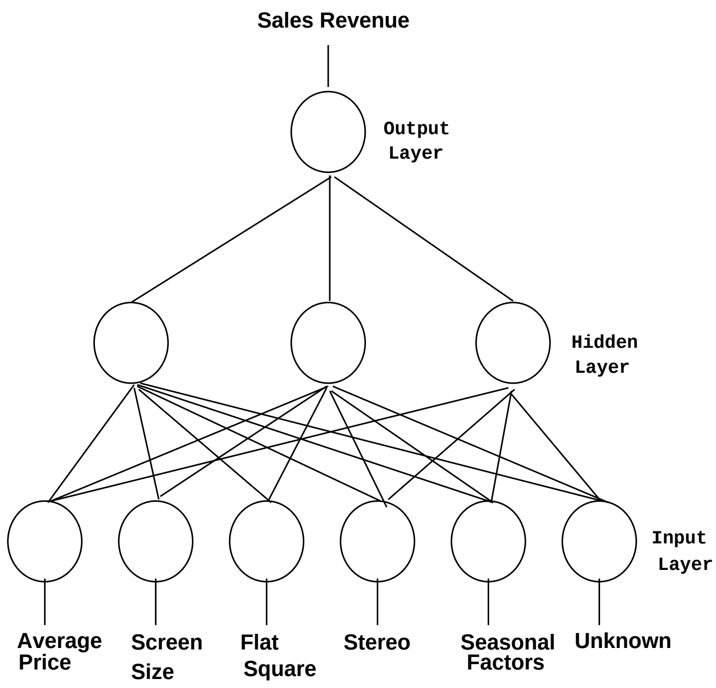

Figure 2: The Final Neural Network Model for CTV Sales in Singapore

is not included. It is quite interesting that the *Stereo* variable is not found to be an influential factor in preliminary tests but later on is found to be among the most influential factors. As stated previously, it is very unusual that *Stereo* is not a influential factor for *Sales Revenue*. This may be due to the fact that much of the information provided by *Stereo* Is already included in the other 17 factors. The final model is one of the simplest in terms of input numbers. The aim of sensitivity analysis is to find the input variable combination which has the most influence on the output variable *Sales Revenue*. Although the result regarding *Stereo* looks contradictory, It really is not. The preliminary testing result shows that if 17 of 18 variables were used to predict *Sales Revenue*, *Stereo* was not influential. The final result shows that if only five or so variables were used *Stereo* was an influential factor on *Sales Revenue*.

There are a few months of the year what are considered to be sales seasons for television sets. These months are mostly festive seasons or school vacations. Inclusion of Seasonal Factors into the final variables further verifies that this is a correct factor to consider. The company might want to consider some special promotions when such seasons arrive. This would serve a two-fold purpose of reminding the company and countering any promotional moves made by competitors.

In terms of the 4-Ps of marketing, three of the final five variables, *Screen Size, Flat-Square* and *Stereo*, are related to **Product**. *Average Price* is related to **Price** and *Seasonal Factor*, is related to **Promotional** time. Since this study focused on Singapore, **Place** was meaningless for the final variables. This information indicates that to the marketing manager, in formulating a marketing mix, might want to look into these main factors further. By doing well in areas related to these factors, sales improvement can be achieved.

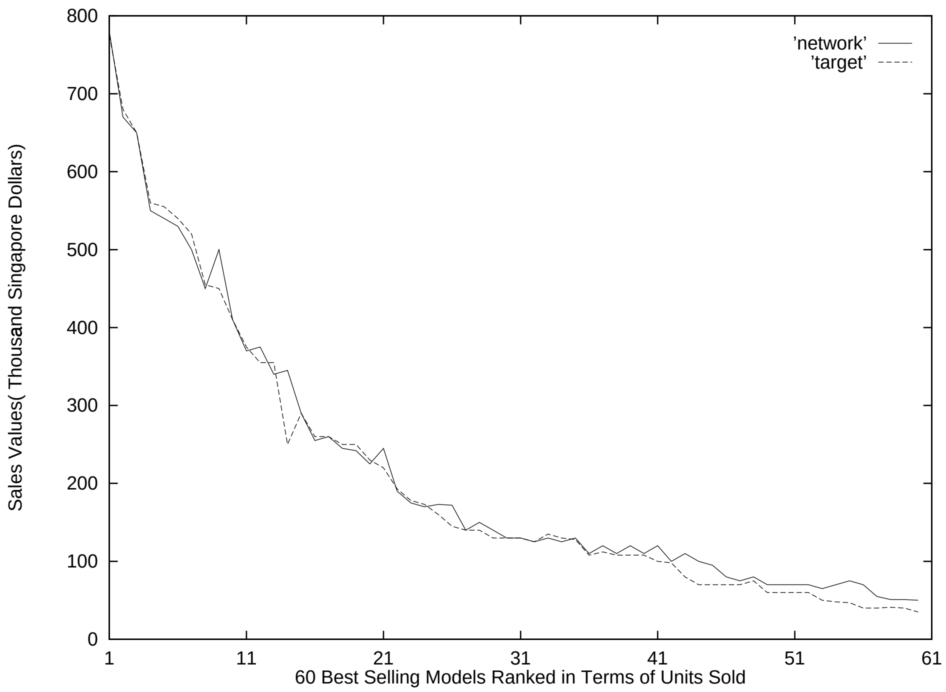

Figure 3: Network Output vs. Target for November 1995 Data

7 Summary and Conclusions

This study has found that, generally, consumers are price sensitive yet desire sets that are not outdated, and that sales fluctuate according to seasons. It has also been demonstrated that the effects of some of the removed variables were overestimated, for example, advertising variables, none of which made it to the final selection. Under the circumstances of this study, it has been found that advertising variables have less effect than do the final selected variables. On the other hand, the influence of some variables on sales revenue has been confirmed by the study, for example, *Seasonal Factors*. Furthermore, the effects of some of the less influential variables has been confirmed, for example, money spent on bus and taxi advertisements. The results obtained are peculiar to Singapore. This study has shed light on the unique brand loyalties of colour television set buyers in Singapore (*Brand Preference* was dropped rather late in the variable reduction process [17]) as well as the particular features that they look for, *Screen Size*, for example.

Through the results of this study, a marketing manager can gain deeper insight into the workings of different marketing variables. This study combined the effects of three groups of variables marketing managers use to make decisions. Knowing the combined effects, the manager need not work with only one group of variables while at the same time speculating on the overall effect when the influence of the other two groups is taken into account. The combined influence of these three groups of variables is not trivial. While preparing and adjusting their marketing plans, or while pricing products, management can pay more attention to the more influential variables found in this study. What is also of interest is that the results of this study can be used immediately.

Implementation of neural networks in marketing and the results obtained by this study have marketing implications. The domain need not be constrained only to the consumer electronics market; in fact, with sufficient data at hand, neural networks can be applied in many other industries. By mean of sensitivity analysis, the number of selected variables was reduced from the original eighteen to five. This study has also shown how these variables can be reduced from networks containing more variables to smaller networks which produced superior results. Interaction between variables that are detrimental to better test results is weakened. *Unknown* was included, and its effect on the network was shown to be important. A trained network with *Unknown* could greatly improve the prediction of neural networks.

The main academic contributions of this study are summarized as follows: (1) implementation of neural networks in analysis of marketing variables; (2) sensitivity analysis and variable reduction through weight analysis; and (3) showing how, by inclusion of *Unknown*, better results can be obtained from neural networks. Apparently, the research method adopted in this study is applicable to other industries and countries as well and need not be confined to the CTV market in Singapore. Further research can focus on complementing the strengths of neural networks with expert and rule-based systems to produce an enhanced DSS. More sensitivity analysis in terms of different variable combinations and comparison with conventional regression models will also be studied.

Acknowledgments

The authors are grateful to the anonymous referees whose insightful comments enabled us to make significant improvements. We would also like to thank Prof Wen-Lian Hsu for his effort devoted to publishing this paper. We thank Yili Li for invaluable discussions and helpful comments.

References

[1] D.C.J. Little, "Decision Support Systems for Marketing Managers", *Journal of Marketing*, Vol.43, Summer 1979.

[2] B.K. Wong, T.A. Bodnovich and Y. Selvi, "A Bibliography of Neural Network Business Applications Research: 1988 - September 1994", *Expert Systems*, August 1995.

[3] H. White, "Connetionist Nonparametric Regression: Multilayer Feedforward Neural Networks Can Learn Arbitrary Mappings", *Neural Networks*, 3, 1990.

[4] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representation by Error Propagation", *Parallel Distributed Processing: Explorations in the Micro-structure of Cognition*, D.E. Rumelhart and J.L. McClelland(Eds.), Volume 1, pp318-362, The MIT Press. 1986.

[5] A. Maren, C. Harston and R. Pap, *Handbook of Neural Network Computing Applications*, Academic Press Inc., 1990.

[6] R.A. Proctor, "Marketing Decision Support Systems: A Role for Neural Networking," *Market Intelligence & Planning*, Vol.10, 1992.

[7] S. Dutta, S. Shekhar, and W.Y. Wong, "Decision Support In Non-Conservative Domains: Generalisation With Neural Networks" , *Decision Support Systems*, 11, 1994.

[8] V. Venugopal and W. Baets, "Neural Networks and Their Applications in Marketing Management", *Journal of Systems Management*, Sept 1994.

[9] S. Schocken and G. Ariav, "Neural Networks for Decision Support: Problems and Opportunities," *Decision Support Systems*, Vol.11, 1994.

[10] T. Hill and W. Ramus, "Neural Network Models for Intelligent Support of Managerial Decision Making", *Decision Support Systems*, Vol.11, 1994.

[11] J.T. Yao, Y.L. Li and C.L. Tan, "Forecasting the Exchange Rates of CHF vs USD Using Neural networks" *Journal of Computational Intelligence in Finance*, Vol.5, No.2., 1997, pp7-13.

[12] P. Kotler, *Marketing Management, Analysis, Planning, Implementation and Control*, Englewood Cliffs, 1994.

[13] R.I. Levin and D.S. Rubin, *Statistics for Management*, Prentice Hall Inc., 1991.

[14] E.M. Azoff, *Neural network time series forecasting of financial markets*, John Wiley & Sons, 1994.

[15] E. Gately, *Neural Networkds for Financial Forecasting*, Jone Wiley & Sons, Inc., 1996.

[16] H.L. Poh, J.T. Yao, and T. Jašić, "Neural Networks for the Analysis and Forecasting of Advertising and Promotion Impact", *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol. 6 No. 4, 1997.

[17] N.Y.W. Teng, *Analysis and Implementation of a Marketing Decision Support System* Honours Year Project Report, DISCS, National University of Singapore, 1996.