



A screening tool based on neural networks

A. M. Cister & N. F. F. Ebecken

COPPE / Federal University of Rio de Janeiro, Brazil

Abstract

Countless words have been written about Customer Relationship Management - CRM. Most of those papers focus on CRM's main strategies -- identification, differentiation, interaction and personalization. These strategies in themselves are perfect, but we should not fail to discuss another crucial topic: what must be done after identifying the customer through a segmentation analysis? In other words, how can we know for sure that our marketing campaigns have been successful in reaching their target -- that is, the customer?

Our goal here is to introduce a tool that will allow us to identify a customer's interests through a segmentation algorithm; help create ideas/prototypes for new products, services, advertising, or promotions; and test marketing actions aimed at specific clusters in order to find out if those clusters have been reached successfully.

This paper presents a model based on a neural network which was created to control the response obtained from customers who have interacted with the company after the campaigns. The main idea is to reduce a lengthy questionnaire by using the principal component analysis -- PCA. The next step will be to use the main variables to classify -- what we call clustering -- and to understand the "new" customer.

The present work was based on a real marketing research (both qualitative and quantitative) conducted by a Telco and Health Company. The idea was to test customers' response following a few marketing actions aimed at specific clusters.

We come to a close by examining the pros and cons of other tools, providing references to other areas of knowledge (such as education, pre-employment tests, and investor's profiling) and presenting our final results.

1 Introduction

The evolution of the market has led to the evolution of marketing techniques. We can safely say that Customer Relationship Management - CRM [1] -- is one of the fastest growing business technology initiatives since the web. These strategies enable us to share a consistent view of front office activities and customer data. Many managers still believe that acquiring customers is better than retaining them.

Are today's companies able to distinguish between the characteristics of a customer they already have from those of a potential one? What is a person's preferred method of paying their bills or seeking advice? Is it through the store itself, a call center, or an Internet advice service? We must remember that new customers do not want to buy the companies' and their managers' preferences and desires. They wish to buy their own preferences. In fact, new customers aren't just buying one's service or merchandise -- they are buying and creating a relationship with the company.

Any organization that can fully understand and explore these issues can become a winner at any kind of market segmentation. Detailed research on customer predilection is critical to pinpoint the most profitable customers, thus identifying those who are no longer worth targeting. This is why customer segmentation plays such a crucial role. The challenge we face is to find the information needed to carry out the segmentation. Of course, we want to start by discovering strategies with marketing actions that target each specific segment.

There is more than one answer to this problem. Information about different customer segments can be obtained through IT systems, database marketing, and data warehouses or through *ad hoc* marketing research in particular. It is important to mention that, after conducting the research, we need to segment customers using some kind of segmentation algorithm. Only then, after performing some statistical analysis (cross validation), are we going to be able to establish a customer's profile.

2 Marketing research

Marketing research plays two important roles in the marketing system. First of all, it is used in intelligence feedback marketing. Marketing research helps decision-makers [2] with data on mixed current marketing perceptions of efficiency and supplies so they can implement the necessary changes. Secondly, it is the main tool used to explore new opportunities in the market. Segmentation research and the research of new products help identify the most profitable opportunities for marketing managers.

According to the American Marketing Association, marketing research can be defined as thus:

The marketing research is the function that connect the consumer, the customer and the public to marketer through information - information used to identify, and to define

marketing opportunities and problems; generate, improve and to evaluate marketing actions; monitor the marketing performance, and to improve the marketing comprehension as a lawsuit. The marketing research specifies the necessary information to board those questions; it formulates the method for the information collection; it manages and it implements the data collection lawsuit; it analyzes the results, and it communicates the discoveries and its implications

In summary, marketing research is used to plan, to collect, and to analyze important data for marketing decision-makers.

3 Factor analysis

Factor analysis (FA) is a multivariate technique used to analyze the interrelationships that exist between a large numbers of variables -- in order to obtain a good overall view of the variance/covariance structure with a minimum loss of information. Its main goal is the reduction of data and its better interpretation [3, 4].

Basically, each factor is grouped by variables tightly correlated to each other -- even though they might show relatively little correlation with other variables belonging to a different group. Factor analysis can be used in numerous kinds of research. FA was the technique used to reduce the number of variables in a survey.

3.1 How to extract the factors?

We can obtain factors through two main techniques:

- Principal component analysis (PCA) - this method provides a unique solution, allowing the original data to be reconstructed from its results. The PCA method evaluates the total variance amongst the variables in such a way that the final result includes as many factors as there are variables.
- The idea behind common factor analysis (FA) makes it very similar to PCA. The difference between the two is that the latter uses the estimate -- the covariance relationship -- to generate the factor solution. Usually, the specific factors can be assumed to be normally distributed. In this case, maximum likelihood estimates the factor loading, so that a specific variance may be obtained.

4 Cluster analysis

The idea in this section is not to teach how to cluster a database, but to review some of the concepts pertaining to clustering techniques. For extensive material on this subject, we suggested Johnson and Wichern [3, pp. 573-609] and Hair *et al.* [5].

Cluster analysis is the name given to a group of techniques employed to identify different groups inside a population, using related variables analysis to characterize each group (customers) as it relates to a given company product, or any other entity. Cluster analysis classifies the variables in each group based on its preferences. As a result, the population is split into a few clusters which exhibit a high degree of homogeneity inside each group, as well as a high degree of heterogeneity between groups.

Currently, many different algorithms may be used to partition a database [6] into similar characteristics or groups. Two of those techniques are discussed here: classic conceptual clustering (hierarchical and nonhierarchical methods) and neural clustering.

For this work in particular, we used a K-means clustering technique. K-means is a nonhierarchical method that enables us to choose the number of groups based on a matrix of distances (similarities). We can start the process by an initial partition of the items into groups, or through an initial set of seeds that will form the center of the clusters (centroid).

5 A case study

A telecommunications company -- which reserves the right not to be mentioned -- ordered the present case study. This company allowed the introduction of an indoor-developed methodology to create a solution to their problem.

The intelligence-marketing team, after a quantitative research, decided to segment their customers in order to identify each customer's profile. Their intention was to promote niche-marketing actions.

In the process, the marketing crew had trouble knowing if their marketing actions were reaching the cluster of interest. The answer was to reduce the size of the questionnaire and to redraw it with fewer variables. This way, whenever a customer got in touch with the company through its call center, the same answered a questionnaire with only a few questions and then, by using specific software, this customer would be classified --on-line -- so the company could verify whether he/she belonged in the cluster of study. This verification was done by using an Excel worksheet which contained the saved NN's weights.

5.1 Methodology

The customers were segmented through the answers given in the survey. The sample group consisted of 3000 customers who were randomly picked with a 1% margin of error and a 95% confidence interval [7].

First of all, we looked for the variable drawback from a factor analysis (principal components) based on a correlation matrix. The Kaiser's criterion method (eigenvalues above 1) was used to determine the number of components (factors) (fig. 1) and the choice of variables was made through communalities analysis (tab. 1)

It was also determined that a variable rotation allowed for a better interpretation of the factors. In this case, the varimax method was chosen, which

attempts to simplify the interpretation by maximizing the variance between the variable loading on each factor, thus simplifying the factors (tab. 2) (tab. 3). It's important to note that all the while the total variance was controlled (tab. 4).

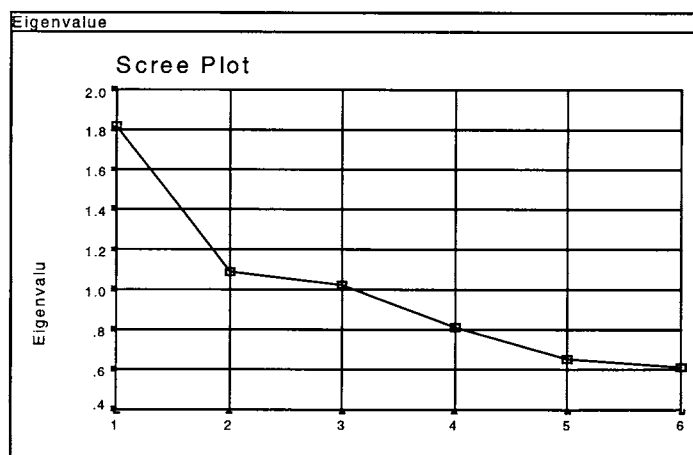


Figure 1: Scree plot: numbers of factors.

Table 1: Communalities analysis.

	Extraction
Frequency Access.	0.759
Facility	0.921
Security	0.634
Price	0.504
Quality	0.548
Recommendation	0.561

After the interpretation of each cluster profile, a neural network was trained with the main variables, that is, the variables which belong to the 3 factors. The main idea was to classify the customers who responded to the marketing campaigns. The marketing crew put forward a campaign aimed at a specific group. The neural network was a perceptron multilayer [8] [9] with a learning rate equal to 0.2, 6 neurons of input and 5 neurons of output, one for each of the previously classified clusters. In a step test, the best neural net showed a 95% accuracy rate.



Table 2: Extraction method: PCA.

Component Matrix			
	Component		
	1	2	3
Quality	0.737		
Recommendation	0.699		
Security	0.669		
Price	0.515		
Frequency access		0.84	
Facility			0.917

Table 3: Rotated component matrix

	Component		
	1	2	3
Recommendation	0.743		
Quality	0.691		
Price	0.679		
Frequency Access		0.85	
Security		-0.617	
Facility			0.959
Extraction Method: Principal Component Analysis.			
Rotation Method: Varimax with Kaiser Normalization.			
Rotation converged in 5 iterations.			

Table 4: Total variance explained.

Component	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.815	30.247	30.247	1.666	27.76	27.76
2	1.09	18.173	48.419	1.21	20.166	47.926
3	1.021	17.02	65.44	1.051	17.513	65.44
Extraction Method: Principal Component Analysis.						

5.2 Screening tool

The screening tool was implemented the following way: After determining the best trained neural net (minimum error), its weights and biases were

implemented in an MS Excel worksheet to be used by the company that hired us. All that is needed now is to go into the next step -- NN -- skipping the unnecessary back-propagation of the error step, in order to classify the customer that answers the "new" questionnaire.

We now need to enter the data pertaining to the new customer's variables that the neural net will use to figure out, on line, the profile within one of the segments (clusters) already defined.

The output of a hidden layer's neuron (h) with a hyperbolic transfer tangent function, with inputs (p) and weights (w) can be defined as:

$$f(h) = \tanh\left(\sum_{i=1}^n w_i p_i + b\right) \quad (1)$$

The output of the neural net (n) can be defined as:

$$f(n) = \sum_{i=1}^n w_i f(h)_i + b \quad (2)$$

The neural net objective function during the training was to minimize:

$$E \left\{ \left| \hat{y} - y \right|^2 \right\} \quad (3)$$

where y = target and $\hat{y}$ = output of the neural network (in other words, minimize the square error). Figure 2 shows the planned outline.

5.3 Pre-test

In order to test and fit this model, we used 60 customers randomly picked from different clusters in the original research database. Calls were made requesting that the customers answer the new questionnaire (redraw). From those 60 customers who answered the "new" questionnaire, 46 customers -- or 76 % of the total number -- had been classified correctly, which means that, in the original cluster 14 customers -- or 23 % of the total -- had been misclassified.

If, at first, this model seems to have yielded average results, having been able to better analyze those 14 clients, it was verified that nine (15 %) of those customers, who no longer belonged in their original cluster, had the biggest distances from the centroid, which is a very interesting fact.

A second analysis must be carried out -- which involves the attitude theory and the customer's habits. We can't forget that this kind of filter is influenced both by the mood of each customer at the moment of the questionnaire and by the behavioral component of an attitude that is reflected by one's tendency to respond a certain way to an objective or activity.

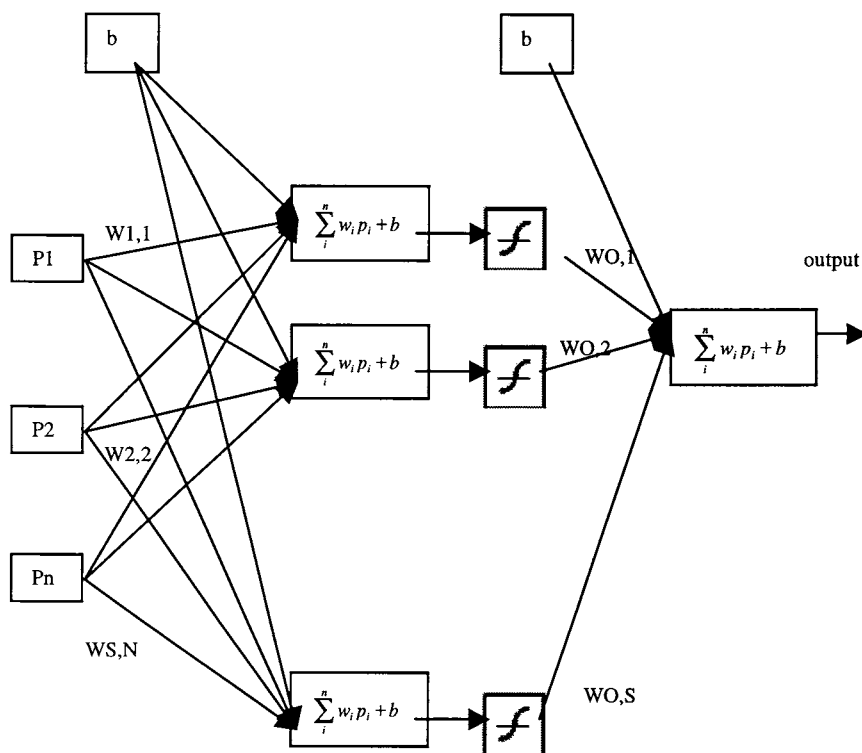


Figure 2: planned outline.

6 Conclusion

The present paper originated in a study of a Telco's needs. At the present time, only the pre-test has been applied. Many people deem the results above as average, but we believe them to be excellent results, as we must keep in mind that our work depends on the moods, attitudes and habits of human beings. Trying to interpret and to put in equation human attitudes is a very challenging task. Our next step will be to use the factors as variables and not the variables that belong to each factor to redraw the questionnaire. Perhaps each factor can work as a profile component; however, we need more tests to validate the model.

On the one hand, it was very stimulating to have found a tool that can measure whether our marketing campaigns have indeed reached our target customers; on the other, we've become conscious of the fact that it is very difficult to control the margin of error since no one really knows how to work with the human personality.

It can be said that the screening tool is relevant to several areas of knowledge, such as education, human resources, financial risks and many other areas.

One must not forget that future events, risks and uncertainties, both individually and collectively, could, to a certain degree, cause actual results to differ substantially from those expressed or implied in these statements. Such factors include changes in worldwide economic conditions, which could considerably impact consumer spending and consumer debt; changes in the demand for the company's products and services; risks associated with the incorporation of acquisitions and other investments; and other factors.

As a learning instrument, we can say that the screening tool allows us to recognize global changes that impact the competitive position of the market share; helps us to focus resources on market opportunities; allows us to identify the company's competitive advantages and core competencies; and finally, it gives us the necessary tools and models for developing a successful strategic marketing plan.

References

- [1] Cister, A M.; Ebecken, N. F. F. - Data Mining III: CRM throught DM: a case study - WIT Press – Ashurst Lodge – U.K. 2002.
- [2] Malhotra, Naresh K., *Marketing Research: An Applied Orientation*, Prentice Hall: Upper Saddle River, New Jersey 07458, 2nd. ed., 1996.
- [3] Johnson, R. A, Wichern, D. W, *Applied Multivariate Statistical Analysis*, Prentice Hall: Upper Saddle River, New Jersey 07458, 3rd ed., pp. 396-448, 1992.
- [4] Johnson, D., *Applied Multivariate Methods for Data Analysts*, Cole Publishing Company, USA, California, 3rd ed., pp. 147-213, 1998.
- [5] Hair, Joseph F. JR. Anderson, R. E.; Tatham, R. L.; Black, W. C. *Multivariate Data Analysis*. Prentice Hall: Upper Saddle River, New Jersey 07458, 4th ed., pp. 469-502, 1998.
- [6] Han, J., Kamber, M., *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, San Diego, California, 3rd ed., pp. 335-389, 2001.
- [7] Kish, L, *Survey Sampling*, John Wiley & Sons: Canada, 1965.
- [8] Haykin, S., *Neural Networks: A Comprehensive Foundation*. Macmillan: New York, pp. 138-220, 1994.
- [9] Hertz, J., Krogh, A. & Palmer, R. G., *Introduction to Theory of Neural Computation*. Addison Wesley, Redwood City, CA, pp. 115-156, 1991.

