ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

MARKET SEGMENTATION AND CLUSTER ANALYSIS

Ankit Jatav*[1] and Akhilesh Salumuri[2]

ABSTRACT

The objective of the research is to consider a self-organizing neural network for segmenting the tourist market. A back-propagation neural network is used to complement the segmentation by generating additional knowledge based on input–output relationship and sensitivity analyses. The findings of the self-organizing neural network indicate three clusters, which are visually confirmed by developing a comparative model based on the test data set.

1. INTRODUCTION

Marketing strategists of destination organizations often encounter the problem of how to appropriately segment the market and package differentiated products and services for target segments. Segmentation is a methodological process of dividing a market into distinct groups that might require separate experiences or marketing service mixes. Customer clustering is one of the most important techniques used to identify these segments.

It includes mainly 2 processes:

a. The basis of this generally includes the identification and assessment of various tourist characteristics (such as demographics, socioeconomic factors, and geo- graphic location).
b. The basis of this generally includes the identification and assessment of various tourist characteristics (such as demographics, socioeconomic factors, and geo- graphic location)

Target marketing is a strategy that aims at grouping a destination's markets into segments so as to aim at one or more of these by developing products and marketing programs tailored to each.

2. OBJECTIVE

The objective of the research is to consider the use of a self-organizing (SOM) neural network for segmenting the international tourist market. A back propagation (BP) neural network (based on the output provided by the former) was also used to complement the process by generating additional market knowledge about the relationship between the inputs used and the macro segments obtained from the application of the SOM model. Input–output relationship and sensitivity analyses were used for the purpose of extracting additional market knowledge about the macro segments.

3. METHODS

a) Elementary Percentiles.
b) Quartiles.
c) Factor Analysis.
d) Principle Components.
e) Cluster Analysis.
f) Current segmentation methodologies and clustering techniques have evolved to include the use of artificial neural networks for segmenting markets.

4. SEGMENTATION TECHNOLOGY

Artificial neural networks are often compared to traditional multivariate techniques (such as multiple linear regression) as a means of benchmarking linearity vs. nonlinearity of a data set. artificial neural networks are superior to linear techniques when compared in terms of accuracy. This is possibly not unusual, as the universal approximation theorem introduced by Hornik, Stinchcombe and White (1989) suggests that a properly trained

ISSN: 2394-7659
IMPACT FACTOR- 2.789

IJERMS

International Journal of Engineering Researches and Management Studies

artificial neural network can never be worse than a linear classifier. Although artificial neural networks also have limitations in respect of explanation, they offer advantages in terms of learning ability, flexibility, adaptation, and knowledge discovery.

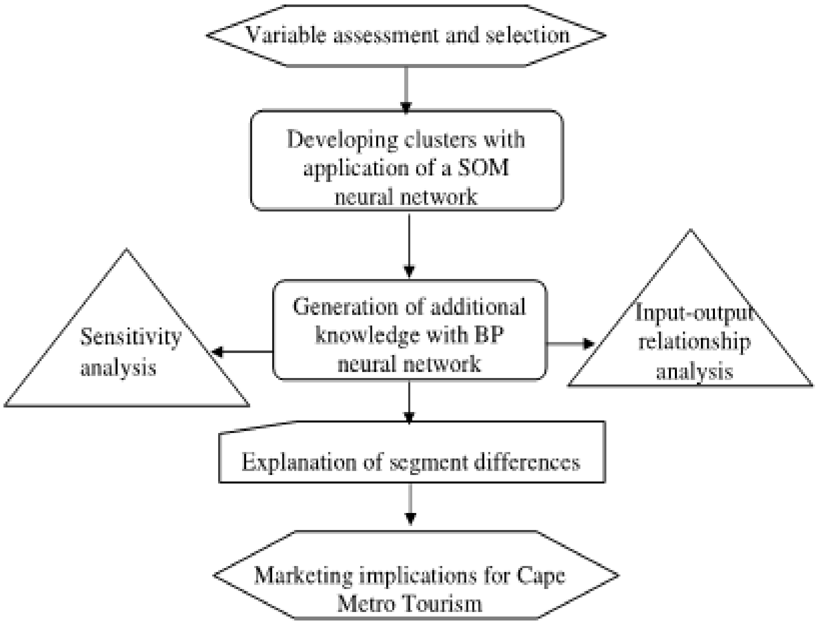

Figure 2. A graphic representation of the research methodology

5. CONCLUSION

The research findings suggest that the methodology and application of a SOM neural network is useful for segmenting the international tourist markets. Furthermore, using a BP neural network model also provides additional market knowledge and, together with the SOM approach for segmenting the tourist market, may improve Cape Metro Tourism's understanding of the international market.

The deployment of the neural network models appears to have merit, even if the intention is only to use the technology as a basis to enhance the market strategists' and media planners' understanding of changing behavior among tourists within and between the macro segments. If this were the rationale for deployment of the neural networks, various micro segments could be identified with an opportunity to disaggregate data and develop macro segments. The use of the segmentation model will assist to achieve these analytical objectives.

6. CLUSTER ANALYSIS

The term cluster analysis itself does not signify a particular method or algorithm. Any assumption to the data distribution is not needed. Clustering groups the data according to related variables. Data can be sort into groups through numerous ways. The choice of method depends mainly on size of data file. Methods commonly used for small data sets are impractical for data files with thousands of cases. SPSS has three different procedures that can be used to cluster data: hierarchical cluster analysis, k-means cluster, and two-step cluster. If your data file is large (even 1,000 cases is large for clustering) or a mixture of continuous and categorical variables, you should use the SPSS two-step procedure. On the other hand, if the data set is small and you want to easily examine solutions with increasing numbers of clusters, you may want to use hierarchical clustering. If the number of clusters is known and you want to cluster a moderately sized data, k-means clustering is a suitable method.

International Journal of Engineering Researches and Management Studies

Hierarchical Clustering

Clusters can be formed through numerous ways. Hierarchical is the most straight forward method. It can be either agglomerative or divisive. In the beginning of Agglomerative clustering, each case is a cluster in itself, and as we proceed, similar cluster keep on merging. The algorithm ends with everybody in one jolly, but useless, cluster. Whereas, Divisive clustering starts with everybody being in one cluster and ends up with everyone in individual clusters. Agglomerative hierarchical clustering doesn't let cases separate from clusters that they've joined. Once in a cluster, always in that cluster. To form clusters using a hierarchical cluster analysis, you must select: A criterion for determining similarity or distance between cases A criterion for determining which clusters are merged at successive steps The number of clusters you need to represent your data 364 There is no right or wrong answer as to how many clusters you need. It depends on what you're going to do with them. To find a good cluster solution, you must look at the characteristics of the clusters at successive steps and decide when you have an interpretable solution or a solution that has a reasonable number of fairly homogeneous clusters.

K-Means Clustering

Hierarchical clustering requires a distance or similarity matrix between all pairs of cases. That's a humongous matrix if you have tens of thousands of cases trapped in your data file. Even today's computers will take pause, as will you, waiting for results. A clustering method that doesn't require computation of all possible distances is k-means clustering. It differs from hierarchical clustering in several ways. You have to know in advance the number of clusters you want. You can't get solutions for a range of cluster numbers unless you rerun the analysis for each different number of clusters. The algorithm repeatedly reassigns cases to clusters, so the same case can move from cluster to cluster during the analysis. In agglomerative hierarchical clustering, on the other hand, cases are added only to existing clusters. They're forever captive in their cluster, with a widening circle of neighbors. The algorithm is called k-means, where k is the number of clusters you want, since a case is assigned to the cluster for which its distance to the cluster mean is the smallest. The action in the algorithm centers around finding the k-means. You start out with an initial set of means and classify cases based on their distances to the centers. Next, you compute the cluster means again, using the cases that are assigned to the cluster; then, you reclassify all cases based on the new set of means. You keep repeating this step until cluster means don't change much between successive steps. Finally, you calculate the means of the clusters once again and assign the cases to their permanent clusters.

Two-Step Cluster

When you have a really large data set or you need a clustering procedure that can rapidly form clusters on the basis of either categorical or continuous data, neither of the previous two procedures fills the bill. Hierarchical clustering requires a matrix of distances between all pairs of cases, and k-means requires shuffling cases in and out of clusters and knowing the number of clusters in advance. The SPSS TwoStep Cluster Analysis procedure was designed for such applications. It requires only one pass of data (which is important for very large data files), and it can produce solutions based on mixtures of continuous and categorical variables and for varying numbers of clusters. The clustering algorithm is based on a distance measure that gives the best results if all variables are independent, continuous variables have a normal distribution, and categorical variables have a multinomial distribution. This is seldom the case in practice, but the algorithm is thought to behave reasonably well when the assumptions are not met. Because cluster analysis does not involve hypothesis testing and calculation of observed significance levels, other than for descriptive follow-up, it's perfectly acceptable to cluster data that may not meet the assumptions for best performance. Only you can determine whether the solution is satisfactory for your needs.

REFERENCES

1. *Bek, M., T. Grosboll-Poulsen, and M. Kristoffersen 2002 Evolutionary Trained KohonenNetworks as Classifiers for Human*
2. *Utterances. Internal Working Paper. Leeds: University of Leeds. Blum, A.*
3. *1992 Neural Networks in C++: An Object-orientated Framework for Building Connectionist Systems. New York: Wiley.*
4. *Crusader Systems 1998 Basic Modelgen Version 1.6. Pretoria: Crusader Systems.*
5. *Deboeck, G. 1994 Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic*

ISSN: 2394-7659
IMPACT FACTOR- 2.789

International Journal of Engineering Researches and Management Studies

6. *Financial Markets. New York: Wiley. De Tienne, K., and L. Lewis*
7. *2003 Artificial Neural Networks for the Management Researcher: The State of the Art. <http://aom.pace.edu/rmd/2003forum/Neural.Nets.pdf>.*
8. *Galloway, G. 2002 Psychographic Segmentation of Park Visitor Markets: Evidence for the Utility of Sensation Seeking. Tourism Management 23:581–596.*