Available online at www.sciencedirect.com

ScienceDirect

Procedia Economics and Finance 27 (2015) 77 – 84

Procedia Economics and Finance

www.elsevier.com/locate/procedia

22nd International Economic Conference – IECS 2015 "Economic Prospects in the Context of Growing Global and Regional Interdependencies", IECS 2015

A Data Mining Based Approach to a Firm's Marketing Channel

Esra Kahya Ozyirmidokuz[a,*], Kumru Uyar[b], Mustafa Hakan Ozyirmidokuz[c]

[a]*Erciyes University, Computer Technologies Department, Melikgazi, Kayseri, 38039, Turkey*
[b]*Nuh Naci Yazgan University,Production and Marketing Department, Kayseri, 38170, Turkey*
[c]*Bosch Thermotechnic, Ankara, 06810, Turkey*

Abstract

Firms need to collect and analyze marketing data in order to have a competitive advantage in the sector. The aim of this research is to extract knowledge from an international firm's marketing channel to improve the efficiency of the marketing system. The Cross Industry Standard Process for Data Mining (CRISP-DM) is used to analyze the survey data. Data are clustered by applying a Kohonen Self Organizing Map (SOM) to reduce the attributes. Anomaly detection analysis is applied. We generate a C5.0 Decision Tree (DT) model used for predicting the marketing channel firms' complaints with very high accuracy. Decision rules are also extracted.

Peer-review under responsibility of Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu"
Keywords: Knowledge discovery in databases; data mining; DT, rule induction, marketing channel complaints

1. Introduction

A marketing channel, which delivers a firm's products and services to consumers, is one of the critical success factors in a marketing system to achieve marketing objectives. If a firm doesn't know its marketing channel well, it

* Corresponding author. Tel.: +90-532-582-8886.
E-mail address: esrakahya@erciyes.edu.tr (E. K. Ozyirmidokuza)

Peer-review under responsibility of Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu"
doi:10.1016/S2212-5671(15)00975-2

can't manage it. Firstly, the firm must collect data from its marketing channel. The firm needs to analyze data seriously to have a competitive advantage in the sector.

Data Mining (DM), which is the process of automatically searching large volumes of data to extract knowledge from them a in a human-understandable structure, helps analysts to recognize relationships within data.

Applying DM techniques to marketing data is extremely useful to find interesting, previously unknown, hidden patterns, which can then be better defined, in massive datasets. In this manner, DM helps to find important knowledge from the marketing channel. The achieved knowledge has a strategic importance in terms of competition and improvement of marketing and production for the firm. This is because knowledge achieved can help to improve the communication between the marketing channel and the firm by better controlling the processes, and by knowing the details about them.

One important type of knowledge that can be obtained from data mining is the decision tree (DT), which is constructed from existing data to classify future data. DTs are an effective method of classifying data set entries and can provide good decision support capabilities. DTs have several advantages over other data mining methods, including being human- interpretable, well-organized, computationally inexpensive, and capable of dealing with noisy data. Due to these merits, DTs are probably the most popular mining method [1]. There have been numerous studies in marketing which use decision trees (DTs) [2, 3, 4, 5, 6, 7, 8, 9].

Among the data mining techniques, cluster analysis helps in the classification of data. Cluster analysis seeks to maximize between-group variances and minimize within-group variances, including both hierarchical and non-hierarchical methods [10].

In the literature, Kohonen's SOM is one of the techniques used for dimension reduction. Malone et al. [11] demonstrated a trained SOM (Self-Organizing Map) which could provide initial information for extracting rules that describe cluster boundaries. Wang et al. [12] used an SOM for pattern analysis and a fuzzy inference system to capture the chaotic trend to provide short-term (hourly) and long-term (daily) Web traffic trend predictions. Fessant et al. [13] used Kohonen SOMs and they showed how the mining of network measurement data can reveal the usage patterns of ADSL customers. Maiorana [14] proposed a feature selection method based on a clustering algorithm belonging to the Kohonen SOM family. Gomez-Carracedo et al. [15] applied Kohonen SOMs to perform pattern recognition in four datasets of roadside soil samples obtained in four sampling seasons over a one year period. They used CART as an objective variable selection step before the SOM grouping. Eshghi et al. [16] compared three clustering techniques: traditional clustering methods, Kohonen maps and latent class models. Nohuddin et al. [17] introduced a technique that uses frequent pattern mining and SOM techniques to identify, group and analyze trends in sequences of time stamped social networks so as to identify interesting trends. In recent years, the Kohonen SOM method has been used in marketing [18, 19, 20, 21, 22].

In this research, we use CRISP-DM, which was developed in 1996 by analysts representing DaimlerChrysler, SPSS, and NCR. CRISP provides a nonproprietary and freely available standard process for fitting DM into the general problem-solving strategy of a business or research unit. According to CRISP-DM, which is shown in Fig. 1, a given DM project has a life cycle consisting of six phases. The phase sequence is adaptive. That is, the next phase in the sequence often depends on the outcomes associated with the preceding phase [23].

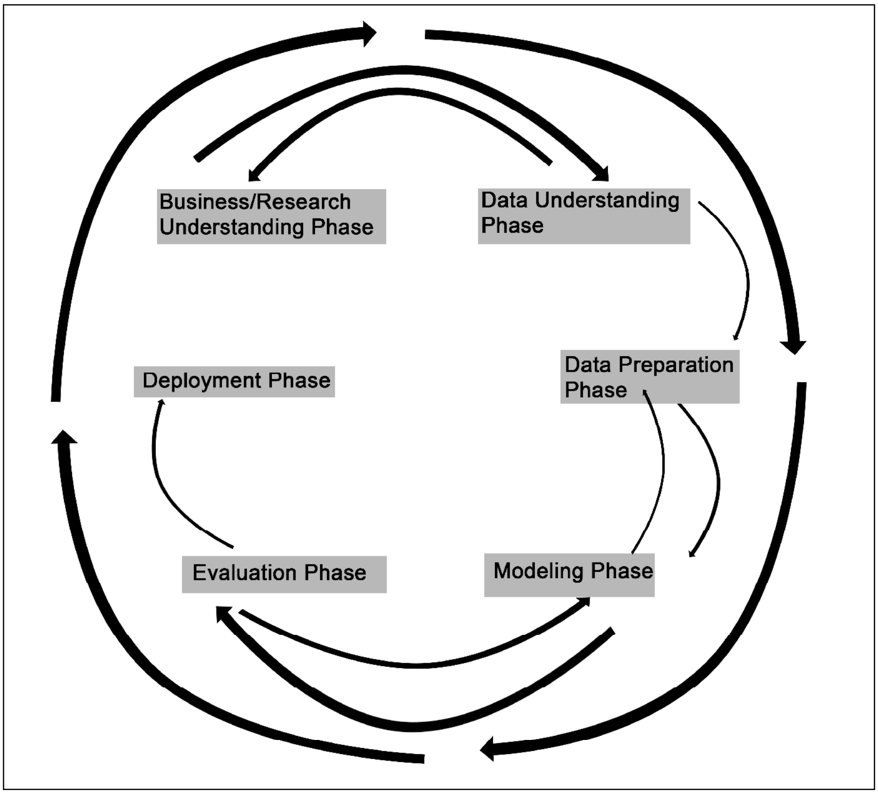

Fig. 1. CRISP-DM process

In the data preparation phase, firstly, we applied anomaly detection analysis which is used to reduce the number of records in a database. In addition, we determined the important features with Kohonen SOMs as a dimension-reduction technique to reduce the features. Although many different techniques could have been used in this study, e.g., PCA, factor analysis,and attribute relevance analysis, we used Kohonen Networks (KNs) in clustering due to the strength of Kohonen maps that lies in their ability to model non-linear relationships between data. The Kohonen map is one of the major unsupervised artificial neural network models. Kohonen Maps are useful tools for DM models with large data sets. High-dimensional data are projected to a lower dimension representation scheme that can be easily understood. In addition, Kohonen Maps can be used to process qualitative variables as well as quantitative ones [24].

DT induction is the modeling step in the forecasting process and consists of the determination of the DTs to generalize previously defined network behavior classes. DTs are easier to understand and they offer an acceptable level of accuracy. Several advantages of the DTs as a classification tool are given in the literature [25]. In this research, a C5.0 DT is applied to the preprocessed data in order to predict the marketing channel firms' complaints about their international firm. Decision rules are extracted. The managers of the firm can easily understand these rules and predict the marketing channel firm's future behavior.

The paper is organized as follows. The Kohonen Network (KN) is introduced in Section 2. Section 3 gives the details of the application. The conclusions are drawn in Section 4.

2. Kohonen Networks

The identification of information, or patterns, in large subsets of data is a property of the fields of data-mining and feature extraction. Unsupervised learning techniques are a subset of these fields which enable the identification and grouping of patterns without having seen that pattern before or having its key characteristics described; to do this, a similarity measure is defined and the groups are clustered together into a lower dimensional space. Self-Organizing Maps (SOM) are one such technique which enable the mapping of data with a large feature set into two-dimensional space. Moreover, SOMs allow the visual understanding of data structures. These can then be used as an aid in identifying and classifying anomalies in the datasets [26].

The goal of SOMs is to convert a complex high-dimensional input signal into a simpler low-dimensional discrete map. Thus, SOMs are very suitable for cluster analysis, when the underlying hidden patterns among records and fields are sought. SOMs structure the output nodes into clusters of nodes, where nodes in closer proximity are more similar to each other than other nodes that are farther apart. A typical SOM architecture is shown in Fig. 2 (Larose, 2005).

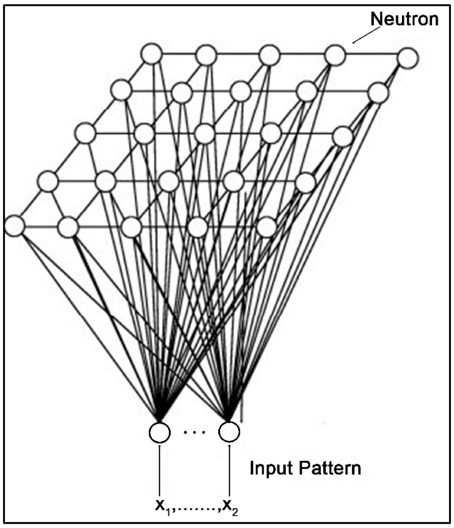

Fig. 2. Topology of a simple SOM

KNs are a type of SOM, which itself represents a special class of neural networks. KNs can be considered as a non-hierarchical method of cluster analysis. As non-hierarchical methods of clustering, they assign an input vector to the nearest cluster, on the basis of a predetermined distance function, but they try to preserve a degree of dependence among the clusters by introducing a distance between them. Consequently, each output neuron has its own neighborhood, which is expressed in terms of a distance matrix. The output neurons are characterized by a distance function between them, described using the configuration of the nodes in a unidimensional or bidimensional space [27].

3. Application

3.1. Understanding Data

Data are collected manually from 300 firms participating in an international firm's marketing channel by one-to-one surveys. The collected surveys are converted to a data matrix which includes 20 features for each firm including education, sex, age, marital status, payment arrangements, the facility preferences provided by the firm, profession, number of trade marks, etc. The sex feature is eliminated manually because all the marketing channel firms' owners, are male except for one, In this phase, the basic statistical analysis and some exploratory graphical analysis techniques are used to understand the data matrix.

Web graphs are performed to understand the relationships between features. We will give two examples. Figure 3 illustrates the relation between educational status and the expectations of the marketing channel firms' owners from the manufacturer firm. We can easily see that the marketing channel firms' owners with a *bachelor's degree* prefer *traveling and training* from the manufacturing firm. However the owner with primary school education does not prefer *traveling and training.* We can also see from Figure 4 that marketing channel firms' owners with lower educational status prefer to communicate with regional representatives.

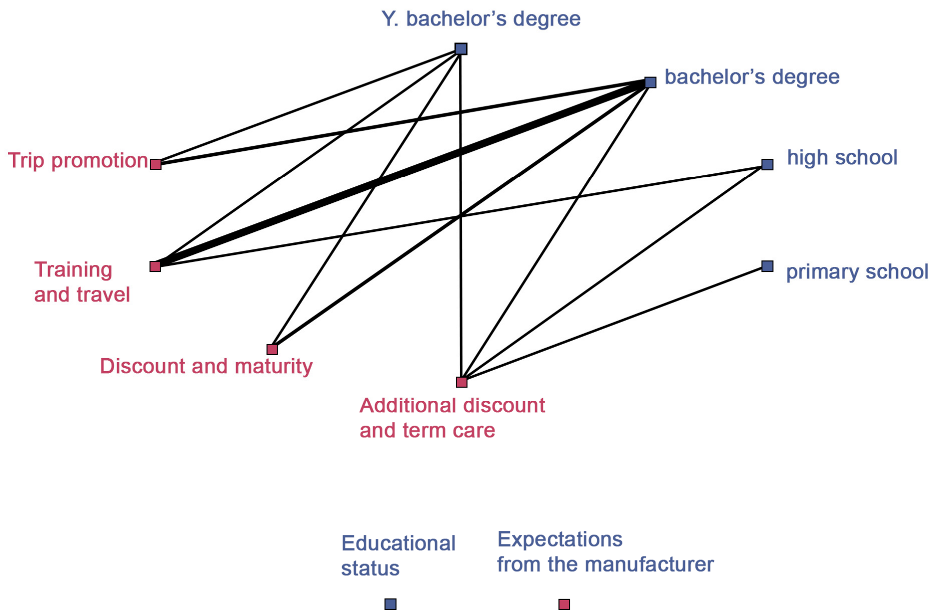

Fig. 3. Web graph of educational status and expectations from the manufacturer

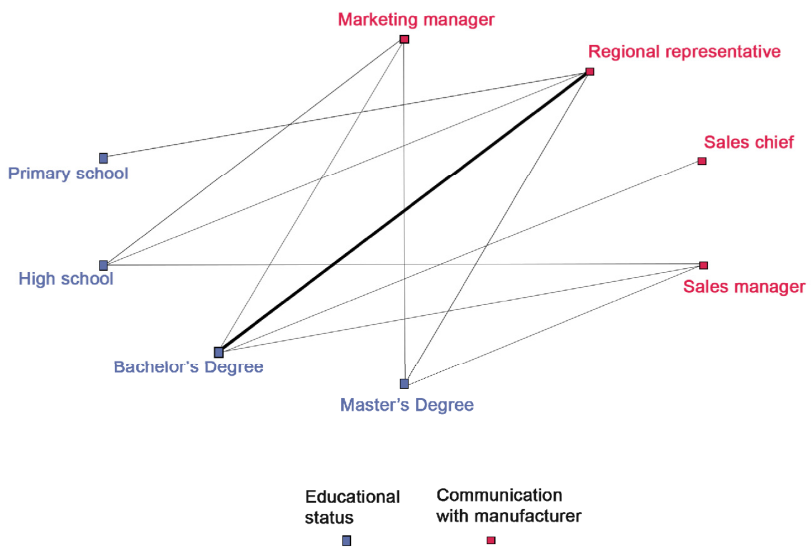

Fig. 4. Web graph of educational status and communication with manufacturer

Although classical statistics work well in giving a picture of the current overall data matrix, they are unable to find hidden knowledge waiting to be discovered. In this manner, DM is used to find previously unknown and useful knowledge.

3.2. Data preparation phase

Anomalous records which show different behavior from the previously measured values in the data matrix must be detected. After we applied anomaly detection analysis, we detected two groups of records, one with 236 and the other with 64 records. After anomaly detection analysis, only three records which were greater than the anomaly index level (1.456) were eliminated from the data matrix.

It is known that feature selection improves model accuracy in the knowledge discovery process. In this research, Kohonen SOM is used to choose a subset of input variables by eliminating features with little predictive information. Records are grouped by KN (Kohonen Network) so that records within a group or cluster tend to be similar to each other, and records in different groups are dissimilar. For the first 20 cycles, the neighborhood size was set at R=2, and the learning rate was set to decay linearly, starting at $\eta = 0.3$. Then, for the next 150 cycles, the neighborhood size was reset to R=1 while the learning rate was allowed to decay linearly from $\eta = 0.3$ to $\eta = 0.1$. The neurons are organized into two layers, namely the input layer and output layer. In the study, the input layer has 45 neurons and the output layer has 12 neurons. The self-organizing map identified 12 different clusters. The clusters and number of records are illustrated in Table1. After applying the Kohonen SOM model, we eliminate 11 features from the data matrix including sex and communication with the manufacturer.

Table 1. The SOM clusters

x	y	Number of records
0	0	42
0	1	22
0	2	48
1	0	22
1	1	4
1	2	14
2	0	15
2	1	6
2	2	26
3	0	47
3	1	2
3	2	49
Total		297

After preprocessing, there are 8 features in the data matrix as follows; *Educational status, Preferences about quality or price, Number of trade marks, Marital status, Profession, Expectations from the manufacturer, Different sectors, Payment arrangements.* These features are the inputs of the DT model. The output of the model is *Complaints about the manufacturer.*

3.3. Decision tree induction

DTs are one of the popular methods that have been used for knowledge discovery in databases. Tree models can be defined as a recursive procedure, through which a set on statistical units are progressively divided into groups, according to a division of an explanatory variable to split and the choice of a splitting rule for such variable, which establishes how to partition the observations. The main result of a tree model is a final partition of the observations. To achieve this it is necessary to specify stopping criteria for the division process[27]. DTs provide an easy to understand overview for users without a DM background with high classification accuracy. They also provide a tree model of the problem and various alternatives in an understandable format without explanation. The acquired knowledge is usually quite understandable and can be easily used to obtain a better understanding of the problem. In addition, DTs assist in making decisions with existing information. They have satisfactory performance even when the training data are highly uncertain [28]. In this study, a C5.0 DT model is run to build a DT for predicting the complaints of the marketing channel's firms. The aim of the classification is to find similar data items which belong to the same class.

Modeling is performed with C5.0, which is one of the popular DT modeling algorithms and an extension of the earlier well known algorithm ID3. A C5.0 DT model which is acceptable with 10 tree depths is achieved. The mean is 91.3 and the standard error is 1.4 for the cross validation. The relations and knowledge are acquired.

One hundred and ninety one if-then rules are generated to express the process in English. The following examples illustrate some of the rules:

- If the *Number of trade marks* > [4] and *Profession* is ["Engineering"] and *Expectations from manufacturer* are ["Additional discount and term care"] then the *firm's complaint* is ["communicating problems"],
- if *Educational status* > ["bachelor's degree"] and *Profession* is ["wholesaler"] and *Expectations from manufacturer* are ["Additional discount and term care"] and *Giro* <= [15000000] then the *firm's complaint* is ["money collecting difficulties"],

Stratified ten-fold cross-validation accuracy evaluation was used to train and test the data matrix. The accuracy rate of the model is 92.67 % which is shown in Table 2.

Table 2. Model accuracy

	Number of records	Percent of records
Correct records	278	92.67%
Wrong records	22	7.33%
Total records	300	100%

4. Conclusions

Today, marketing channel decisions are as important as the decisions companies make about the features and prices of products [29]. In this research, we applied a DM framework and we presented a decision tree induction from marketing channel data to improve the efficiency of the marketing system. DM techniques were implemented to marketing survey data. We explored the use of different pre-processing and DM techniques including anomaly detection analysis, Kohonen SOM, and C5.0 DTs. This research included attribute reduction using KNs. A C5.0 DT which was used for the classification of the data set with 10 tree depths was generated. The accuracy rate of the model was 92.67%. The DT model lays out the data matrix clearly so that all options can be explored. This acquired knowledge may be used to predict the future behaviors of the marketing channel firms. Data are processed into a manageable format, and decision rules are also generated. The DT model helps managers to understand the marketing channel firms. This research is also important to assess the future complaints of the firms and to plan future marketing developments. If we evaluate the current marketing channel and plan for capacity needs we will achieve a better marketing system performance. Consequently, the knowledge obtained will improve the performance of the marketing system.

Alternative DM techniques using artificial intelligence methods can be studied in future research to compare various approaches and to implement this framework.

Acknowledgements

This work was supported by Erciyes University Research Fund, Project Number FBA-2014-5364, Turkey.

This project is also supported by Bournemouth University, Faculty of Science and Technology, Software Systems Engineering, UK.

References

Chen, Y-L., Hu, H-W., Tang, K., 2009. Constructing a Decision Tree from Data with Hierarchical Class Labels. Expert Systems with Applications 36, 4838–4847.

Crone, S.F., Lessmann, S., Stahlbock, R., 2006. The Impact of Preprocessing on Data Mining: An Evaluation of Classifier Sensitivity in Direct Marketing. European Journal of Operational Research 173, 781–800.

Lee, S., Lee, S., Park, Y., 2007. A Prediction Model for Success of Services in E-Commerce Using Decision Tree: E-customer's Attitude Towards Online Service. Expert Systems with Applications 33, 572–581.

Yada, K., Ip, E., Katoh, N., 2007. Is this Brand Ephemeral? A Multivariate Tree-Based Decision Analysis of New Product Sustainability. Decision Support Systems 44, 223–234.

Kim, S.S., Timothy, D.J., Hwang, J., 2011. Understanding Japanese Tourists' Shopping Preferences Using the DT Analysis Method. Tourism Management 3, 544–554.

Olson, D.L., Chae, B.K., 2012. Direct Marketing Decision Support Through Predictive Customer Response Modeling. Decision Support Systems 54, 443–451.

Duchessi, P., Lauría, E.J.M., 2013. Decision Tree Models for Profiling Ski Resorts' Promotional and Advertising Strategies and the Impact on Sales. Expert Systems with Applications 40, 5822–5829.

Lei, N., Moon, S.K., 2015. A Decision Support System for Market-Driven Product Positioning and Design. Decision Support Systems 69, 82–91.

You, Z., Si, Y-W., Zhang, D., Zeng, XX., Leung, S.C.H., Li, T., 2015. A Decision-Making Framework for Precision Marketing. Expert Systems with Applications 42, 3357–3367.

Berson, A., Smith, S., Thearling, K. Building Data Mining Applications for CRM. New York: McGraw-Hill; 2001.

Malone, J., McGarry, K., Wermter, S., Bowerman, C., 2005. Dm Using Rule Extraction from Kohonen Self-Organizing Maps. Neural Comput and Applic 15, 9–17.

Wang, X., Abrahamb, A., Smith, K.A., 2005. Intelligent Web Traffic Mining and Analysis. Journal of Network and Computer Applications 28, 147–165.

Fessant, F., François, J., Clerot, F., 2007. Characterizing ADSL Customer Behaviours by Network Traffic Data-Mining. Annales des Télécommunications 62 (3–4), 350-368.

Maiorana, F., 2009. Feature Selection with Kohonen Self Organizing Classification Algorithm. World Academy of Science, Engineering and Technology 27, 998–1003.

Gomez-Carracedo, M. P., Andrade, J.M., Carrera, G.V.S.M., Aires-de-Sousa, J., Carlosena, A., Prada, D., 2010. Combining Kohonen Neural Networks and Variable Selection by Classification Trees to Cluster Road Soil Samples. Chemometrics and Intelligent Laboratory Systems 102, 20-34.

Eshghi, A., Haughton, D., Regrand, P., Skaletsky, M., Woolford, S., 2011. Identifying Groups: A Comparison of Methodologies. Journal of Data Science 9, 271–291.

Nohuddin, P.N.E., Coenen, F., Christley, R., Setzkorn, C., Patel, Y., Williams, S., 2012. Finding ''Interesting'' Trends in Social Networks Using Frequent Pattern Mining and SOMs. Knowledge-Based Systems 29, 104–113.

Gómez-Pérez, G., Martín-Guerrero, J.D., Soria-Olivas, E., Balaguer-Ballester, E., Palomares, A., Casariego, N., 2009. Assigning Discounts in a Marketing Campaign by Using Reinforcement Learning and Neural Networks. Expert Systems with Applications 36, 8022–8031.

Lee, J.H., Park, S.C., 2005. Intelligent Profitable Customers Segmentation System Based on Business Intelligence Tools. Expert Systems with Applications 29, 145–152.

Mostafa, M.M., El-Masry, A.A., 2013. Citizens as Consumers: Profiling E-Government Services' Users in Egypt Via Data Mining Techniques. International Journal of Information Management 33, 627– 641.

Dragomir, O.E., Dragomir, F., Radulescu, M., 2014. Matlab Application of Kohonen Self- Organizing Map to Classify Consumers' Load Profiles, 2nd International Conference on Information Technology and Quantitative Management, ITQM, Procedia Computer Science 31, 474 – 479.

Cuadros, J., Domínguez, V.E., 2014. Customer Segmentation Model Based on Value Generation for Marketing Strategies Formulation Alvaro. Estudios Gerenciales 30, 25–30.

Larose, D.T. Discovering Knowledge in Data: An Introduction to DM. Wiley, USA; 2005.

Tuffery, S. DM and Statistics for Decision Making. USA: Wiley; 2011.

Rokach, L. and Maimon, O. Data Mining with Decision Trees. River Edge, NJ: World Scientific; 2007.

Palomoa, E. J., North, J., Elizondo, D. , Luquea, R.M., Watson, T., 2012. Application of Growing Hierarchical SOM for Visualization of Network Forensics Traffic Data. Neural Networks 32, 275–284.

Giudici, P. Applied DM. England: Wiley; 2003.

Kahya-Özyirmidokuz, E., Gezer, A. and Ciflikli, C. , 2012. Characterization of Network Traffic Data: A Data Preprocessing and Data Mining Application. Data Analytics 2012: The First International Conference on Data Analytics, Barcelona, Iaria Think Mind, 18-23.

Hitt, M., Black, S., Porter, L. Management. 2nd ed. Upper Saddle River NJ: Prentice Hal; 2009.