A LOW-RISK EXPERIMENTS FOR CUSTOMER SEGMENTS OF E-MAIL MARKETING CAMPGAINS: A BUSINESS INTELLIGENCE APPROACH

[1]GEORGE SAMMOUR, [2]BENOIT DEPAIRE, [3]KOEN VANHOOF

[1]Management Information Systems Department, Princess Sumaya University for Technology, Jordan
[2,3]Faculty of Applied Economic, Hasselt University, Belgium
E-mail: [1]George.sammour@psut.edu.jo, [2]benoit.depaire@uhasselt.be, [3]koen.vanhoof@uhasselt.be

Abstract- Research in e-mail marketing is divided into two broad areas spam and improving response rate. In this paper a methodology is proposed which allows companies to experiment with their e-mail campaigns to increase the campaigns' response rate. This methodology is particularly suited for companies that are reluctant to experiment with their customer's data which may lead to a drop of the response rate due to unsuccessful changes of the e-mail campaign. The objectives of this research have been achieved in applying a two-step approach. Firstly, homogeneous groups of customers are identified, eliminating largely any hindering heterogeneity. Secondly, customers that are not clicking and/or having a low click rate within their homogeneous groups are identified. The results are promising such that it allowed for making informed recommendations for low-risk experiments on customers having a non/low-click behaviour on a weekly newsletter e-mail.

Keywords- E-mail marketing campaign, response rate, click rate, cluster analysis, decision tree modelling

I. INTRODUCTION

The benefits of an active marketer are described in considerable detail in the customer relationship management (CRM) literature, which suggests that marketers could enhance customer loyalty by being active and in regular contact with their customers. Email offers a promising tool that helps marketers keep in touch with their customers on a regular basis at low cost. The advent of email technology has created a new channel for marketing, as the transition from the old world of business management to the new world of e-customers is unique. The effect of this change in the communication chain is altering the way companies market to their customers and how they deliver their messages and products.

E-mail marketing is a method that affords means of communication between the enterprise and the customer via e-mail. The e-mail sent to the customer contains announcements and advertisements or any declaration of informative content to enhance customer loyalty by being active and in regular contact with the customer. E-mail marketing has become an indispensable form of communication due to the low cost and the transition from the old world of business management to the new world of e-customers. In addition to the ability to track each e-mail sent and grasps data about click rate and response rate which are two main metrics of the activeness of the e-mails sent.

Although practitioners and academics have identified key success factor and key barriers to the development of an effective e-mail campaign, few have attempted to apply existing theories and models. Although e-mail marketing studies have been conducted either by online surveys, by in-depth interviews, by controlled experiments or by tracking behaviour patterns such as click-through links and the visiting patterns, few research have investigated the effects of e-mail characteristics on consumer attitudes and behavioural intentions. Click rate and response rate are two terms couples with e-mail marketing, they indicate the number of clicks in HTML and text-only e-mails. It provides a way to estimate the customer attachment with an e-mail.

II. RELATED WORK

Most mass media venues do not allow marketers to target consumers with a high degree of precision, even though targeting and segmenting are maybe marketing centrepieces (Krishnamurthy, 2001). With this difficulty existing, one of the most recent direct marketing approaches focuses on consumers' preferences and developing a meaningful interactive dialogue (Kent & Brandal, 2003).

Godin (1999), introduced a technique called permission marketing, which seeks permission in advance from consumers to send marketing communications. Consumers provide interested marketers with information about the types of advertising messages they would like to receive.

Marketers then use this information to target advertisements and promotions. The aim is to initiate, sustain and develop a dialogue with customers, building trust and over time stimulating the levels of permission, making it a more valuable asset (Kent & Brandal, 2003). Permission marketing has three specific characteristics that set it apart from traditional direct marketing (Godin, 2000):

- Anticipation: Customers who permit their names to be included on direct-mail lists can anticipate receiving commercial messages.
- Personalization: The sending company can personalize those messages.

- Relevance: The messages will be more relevant to the customers' needs.

These characteristics are what allow marketers to cut through the clutter and speak to prospects as friends. This personalised, anticipated, frequent, and relevant communication has a greater impact than a random message displayed in a random place at a random moment.

There exists some research which builds models to improve response rate by using individual preferences to personalise e-mail newsletters through collecting and analysing such information. Marketing campaigns and products can be customised to appeal better to groups of customers, or the individual. Recent studies look specifically at e-mail communication. For example, a model of online clicking behavior by Ansari and Mela, attempts to predict and improve response rates for e-mail communications (Ansari and Mela, 2003).

E-mail is considered as the highest Return on Investment (ROI) bringing elements of all direct marketing channels, thus, it should have its own strategy and resources. Whether that resource is a team of experts or a single person, the skills required are: "creative design, content management, copywriting, coding, list and database management, strategy, and analysis" (Jenkins, 2009). Successful e-mail marketing requires that a purpose is set for each campaign that is going to be sent. Therefore, a target or a goal for a campaign is determined, and then the campaign results are compared with whatever goal was set. For example, the goal of a promotional campaign may be to increase sales of a certain product or service; or the goal of another campaign may be to simply gather responses to a survey as much as possible.

Bodnar and Cohen (2011), discussed that the campaign must offer something valuable to the recipient on which actions can be made. Whether that value proposition is important information, a link to a discount page, or a specific promotion guiding the customer toward a physical store, depends on the purpose of the campaign itself. If the campaign is not relevant, and lacks a value proposition, the subscriber will most likely disregard that campaign, leading to unwanted results.

Ryan & Jones (2012), argues that it is difficult to define why any particular campaign was better than another if different elements aren't tested. On the other hand, MailChimp (2015) offer a service referred to as A/B testing, which is a service that allows recipient lists to be split into groups based on the elements one wishes to test. The software splits a predetermined number of recipients of the original list into two groups (A and B) compares the results, and then automatically sends the winning version to the rest of the recipient list.

Split testing is used to determine factors that affect the performance of e-mail campaigns. Usually elements that are tested include; subject lines; timing; graphics and structure; call-to-action placement; contents, and offers (MailChimp, 2015)

III. METHODOLOGY

The aim of this research is to identify and improve response rate of e-mail campaigns. This goal is achieved by identifying Homogeneous groups of customers which are not/low responding to e-mail campaigns. Because of their current low response level, these groups of customers have a high potential to increase the overall response rate. Moreover, experimenting with such low response rate groups has a low risk of decreasing the response rate. Consequently, if experimenting fails will not lead to decrease in response rate, since these groups of customers are (almost) not responding. The identification of homogeneous groups of customer is achieved by applying data mining techniques in two phases as follows:

- Identify homogeneous groups of customers based on socio-demographic or other type of customer information.
- Segment customers within each Homogeneous group based on their response/open/click rates (i.e. behavior).

Both phases are accomplished through the use of data mining techniques. Data mining can be defined as the nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in large amounts of data (Fayyad et al, 1996). Depending on the objective of the research, two major categories of data mining can be recognised predictive and descriptive techniques. Predictive data mining techniques are usually applied to problems with the goal to identify a statistical model or set of models that will be used to predict some response of interest. Other data mining problems are exploratory in nature (e.g. to identify clusters or segments of customers), in which case descriptive data mining is applied.

The first phases of finding Homogeneous groups of customer we opted for is the descriptive data mining technique of cluster analysis. This technique seeks to separate data elements into groups or clusters with similar characteristics, such that both homogeneity of elements within clusters and the heterogeneity between clusters are maximised (Hair et al, 1998). This step is important because heterogeneity can hide real effects: applying changes to marketing campaigns for a heterogeneous group of customers might work for some part while be detrimental to another part resulting in a zero net result.

In the second phase each customer cluster will be experimented to further identify non-clicking or low-clicking customers. This will be achieved using Decision Trees (DT), which are mainly used for prediction and classification of unknown cases

problems. In the next two subsections, clustering techniques and DTs will be discussed in detail.

3.1 Clustering Methods

Cluster analysis has been applied in a wide variety of fields, ranging from engineering, computer sciences (web mining, spatial database analysis, textual document collection, image recognition and segmentation), life and medical sciences, to earth sciences, social sciences and economics (marketing, business, CRM) (Everitt et al, 2001; Green, 2004; Arabie & Hubert, 1994; Moustaki and Papageorgiou, 2005; Jiang et al., 2004). According to Fraley and Raftery (2002) cluster analysis is based on heuristics that try to maximise the similarity between in-cluster elements and the dissimilarity between inter-cluster elements. These similarity-based clustering techniques use a specific distance function for elements with qualitative features. For elements consisting of both continuous and qualitative features, a mapping into the interval (0,1) can be applied such that a distance measure can be used. Among the similarity-based techniques, two major approaches can be detected, namely the hierarchical approach and the partitional approach (i.e. K-means). Extensive research has been done in this field of heuristic-based cluster analysis, but the statistical properties of these methods are generally unknown (Fraley and Raftery, 2002), whereas the statistical properties of a second type of cluster analysis, i.e. probability model based clustering techniques (Bock, 1996; Fraley and Raftery, 2002) are better understood.

Following the maximum likelihood approach, the unknown parameter vector is often estimated by means of the expectation-maximisation algorithm. Outliers are handled by adding one or more classes, representing a different multivariate distribution for outliers (Fraley and Raftery, 2002). Typically, if a small cluster appears which is hard to profile by means of the cluster-dependent distributions, one has found a group of outliers. Other possibilities to handle outliers are by using exploratory analysis in advance (Moustaki and Papageorgiou, 2005).

3.2 Decision Trees

Decision tree learning is a method for approximating discrete-valued target functions, in which the learned function is represented by a decision tree. It performs many tests and then tries to arrive to the best sequence for predicting the target. Each test creates branches that lead to more tests, until testing terminates in a leaf node. The path from the root to the target leaf is the rule that classifies the target. The rules are expressed in if-then form (J. Quinlan, 1992). Decision trees have obvious value as both predictive and descriptive models. The training process that creates the decision tree is called induction and requires a small number of passes through the training set. Most decision tree algorithms go through two phases: a tree growing (splitting) phase followed by pruning phase. Given the properties and nature of classification of decision tree algorithms and the nature our data, as will be discussed in the next section, we decided to use the C4.5 decision tree algorithm. C4.5 is not restricted to binary splits and it produces a tree of more variable shape. C4.5 algorithm uses the fact that each attribute of the data can be used to make a decision that splits the data into smaller subsets.

It should be noted that decision trees are mainly used for classification of unknown cases, but in the scope of this research we used DT as a descriptive method and segmentation technique. DT will segment the set of known customers into groups with similar values for the class variable, which will be any response-related criteria (i.e. Click criteria). Furthermore, due to our exploratory use of DT, we are less interested in the generalisation power of the learned model. The DT model will merely allow us to identify once again homogeneous groups of customers with a low response level to e-mail campaigns.

IV. DATA DESCRIPTION

The data collected contains information on 32 weekly electronic newsletters during the period from June 2007 until the end of January 2008, from a customer of Ideaxis that is using the ADDEMAR® platform. The content of the newsletters is divided on the basis of six areas of interest; these areas of interest are wine, Recipes, new products, promotions, health & bio- products and member cards. The layout of the newsletter is depicted in Fig. 1, as shown on top of the newsletter the six areas of interest are listed and for each consumer only the areas he has chosen will be enabled. On registration, subscribers can choose the relevant areas of interest.

Fig.1: campaign newsletter layout.

The content of the newsletter is automatically personalized for each recipient. Also, it is possible for consumers to choose the format of the newsletter so the subscriber has the choice of a simple text email or an HTML email. The downside to text emails is that they are not measurable in terms of open rate (Walrave, 2004), so the open rate will not be considered in this study with regards to customers. The newsletters are prepared in two languages, Dutch and French, which are the official languages in Belgium.

The number of contacts is 31,385 whose 19,609 of them is Dutch-speaking (NL) and 11,776 are French-speaking (FR) customers. In the scope of this study only the Dutch speaking customers are studied for the sake of homogeneity in the data, and that the Dutch speaking customers are almost 63% of the overall contacts, furthermore after analyzing those customers we found out that not all of them received the same number of newsletters since some consumers subscribed late, so we filtered out customers who received all 32 campaigns, which result in a 1172 customers (n=1172). For each customer we collected information such as, gender. email format, interests, number of interests he chose to receive, total emails opened and total emails clicked, after that we calculated the click rate and open rate for each customer, furthermore, for segmentation purposes, we categorized the click rates to non-click, low-click and high-click rates. Finally, the time the email was first opened was also included in the dataset.

V. EXPERIMENTS AND RESULTS

As stated in our problem statement, the focus of this study is to identify homogenous segments of customers which are not responding and/or having a low-click profile to the email newsletters.

5.1 Cluster Analysis

As outlined in the methodology, in the first phase a cluster analysis to remove big parts of heterogeneity in the data is conducted. Latent Cluster Analysis is performed by means of the LatentGold® software version 2.0.9. The values of BIC, AIC and CAIC are used to assess the optimal number of clusters. These statistical Fig.s measure the model fit (a lower score is better).

For clustering customers into homogeneous groups, customers' interests are used as indicators or attributes for clustering. The analysis is run to produce a minimum of two clusters and maximum of six clusters. Table 1 summarises the results revealing the values of BIC, AIC and CAIC. The results show that the values of BIC, AIC and CAIC first goes down when adding more clusters, but at a certain point (4- cluster model) starts to increase. For all three statistics, the minimum is reached at the 3-cluster model. Hence, as the values of BIC, AIC and CAIC suggest, the 3-cluster model is the best model.

Fig. 2 depicts the profile for each cluster. Each line represents one cluster. The horizontal axis represents the six preferences (interests) included in the weekly newsletter. The vertical axis scales the probability that customers belonging to each cluster with specific interest. To define each cluster, the 50% cut-off rule is used. As shown in Fig. 2 the first cluster or group of customers are interested in newsletters related to recipes, the second group are interested to newsletters with topics about all six topics, and the third group are interested in all topics except promotions and member cards. Furthermore, the 3-cluster model has the best trade-off between model complexity and model fit.

Table 1: Cluster analysis results comparing BIC, AIC and CAIC values

Model	L^2 (or LL)	BIC	AIC	CAIC
2-cluster	236.874	-116.406	136.874	-166.406
3-cluster	121.673	-182.148	35.6732	-225.148
4-cluster	121.657	-132.704	49.6579	-168.704
5-cluster	28.4444	-176.458	1989.76	2195.99
6-cluster	18.6379	-136.805	1993.95	2242.64

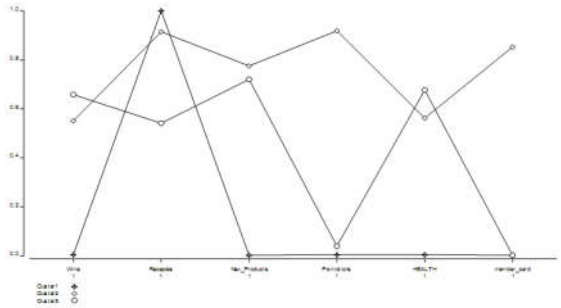

Fig. 2: customer distribution according to interests.

Table 2 summarises the distribution of customers across all clusters with extra statistical information such as e-mail format (TEXT or HTML), click rate and a short description of customers' interests. The majority (87%) of customers belongs to cluster-1, most of them prefer a text-formatted e-mail, while customers in cluster-2 and cluster-3 prefer HTML formatted e-mails.

Table 2: Statistical information of customers in Clusters

Cluster No.	Description	percent	HTML	TEXT	Click Rate
1	Customer interested in recipes	87%	47%	53%	6.1%
2	Customers interested in all 6 categories	8%	67%	33%	16.7%
3	Customers not interested in promotions and member cards	5%	74%	26%	9.9%

Additionally, Table 2 reveals that customers in cluster-1 have lower click rate than customers of the other two clusters. The reason that customers of cluster-1 have a low click rate can be explained by the fact that majority of customers are interested in a single topic (recipes) with text-formatted e-mails (i.e. no hyperlinks). Customers in cluster-2 recorded the highest click rate (16.7%), with preference of receiving HTML-formatted e-mails and all categories. Subsequently, customers in custer-3 preferred HTML-formatted e-mails and click rate of almost 10% with no interests in promotions and member cards.

The results presented so far reveals that the majority of customers (cluster-1) have a low click rate. Such

results already provide useful information with regards to their current e-mail marketing campaigns. Thus, the company must focus on the content of the recipes category, such as including more interactive content (images, videos ...etc.).

5.2 Decision Tree Analysis

The second step of our methodology, performs a decision tree analysis for each cluster in order to find homogenous segments of customers with a low/non clicking profile. To this end, we categorized the click rate into three categories, i.e. non-click, low-click and high click and we will use this categorized variable as the DT class variable. We categorized the click rate as mentioned above into three categories according to the values calculated in the data set for customers, customers who have click rate evaluated to zero have a non-click lick criteria, customers who have a click rate more less than 10% are categorized as low-click, and finally customers having a click rate more than 10% are categorized as having a high-click profile. Besides, having the click criteria as the class variable, we used the gender, email format, interests, and the period of time the customer opened the emails variables as attributes to build the decision trees.

Fig. 3 depicts the DT model for customers of cluster-1 (customers interested in recipes and in text-formatted e-mails). The first number at leaf nodes of the DT represents the correct cases belonging to this branch, while second number represents the error. Recall that customers belonging to this cluster have a click rate of 6.1% (which can be categorized as low-click). However, the DT model reveals that customers who prefer HTML-formatted e-mails (47%) have a high-click rate on recipes. Furthermore, investigating the DT model, one can note that among the 543 customer who preferred text-formatted e-mail, 488 (82%) are non-clicking. On the other hand, 55 (18%) of customers have a low-click behaviour.

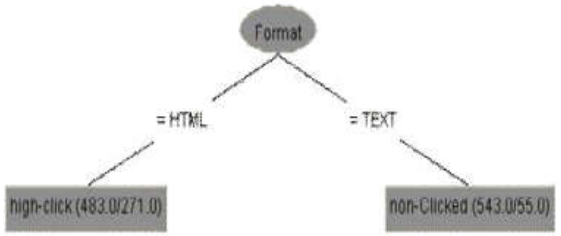

Fig. 3: Decision Tree for Cluster 1

This segment of customers is a perfect group to experiment with. In the worst case you could turn 55 customers from low-clicking into non-clicking, but in the best case, you could turn 488 customers into low or even high clicking customers.

Fig. 4 shows the decision tree for customers belonging to cluster 2, which forms almost 8% of all customers. The Fig. shows two customer segments who are also good candidates for experimenting with. Firstly, there is the group of customers who preferred TEXT e-mails. This group of 30 people have 20 customers which are simply not responding to emails, while the other 10 have low-click behaviour. Secondly, there is a group of customers who prefer HTML e-mails and have no interest in information about new products or wine. This group of 10 customers contains seven low clicking customers are interested in receiving newsletters related to recipes identified by clustering.

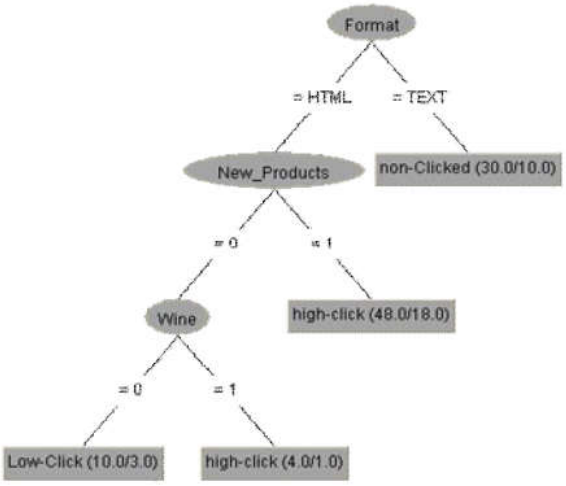

Fig. 4: Decision Tree for cluster 2

Finally, for cluster 3 (Fig. 5), there are three candidate groups for experimenting. In contrast with the previous two clusters, we can now identify two different groups among the customers which preferred TEXT e-mails. Among these customers, we can distinguish between those who prefer information about new products and those who do not. The first group contains 10 customers, who are all non-clicking.

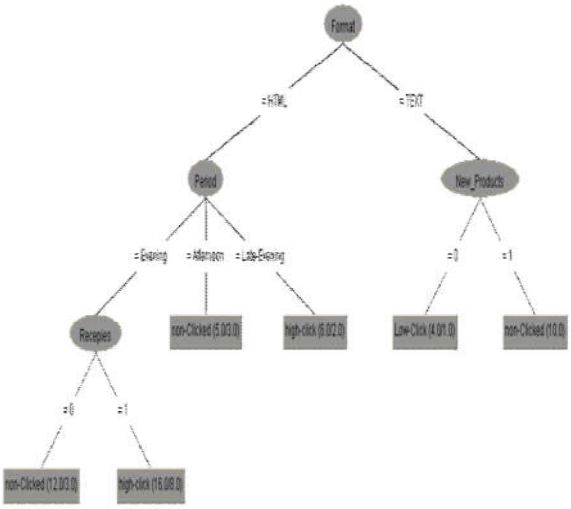

Fig. 5: Decision Tree for cluster 3

This group can be further investigated as one cannot decrease the overall click rate by experimenting. The other group of customers who are not interested in information about new products, contains four customers from which three are low clicking. Note

that this might be considered a too small group for experimenting. One could decide though to group them with customers who are interested in new products. Furthermore, among the customers who are receiving HTML e-mails, an interesting experiment group are those who receive the newsletters in the afternoon and are not interested in Recipes. This group contains 12 customers among which nine are not clicking any links inside the newsletters.

Furthermore, for customers receiving HTML emails, only 10 customers that are not interested in receiving newsletters related to promotions nor Recipes are not clicking, while only 5 customers interested in promotions and not interested in Recipes nor new products have low-click rate. What is interesting in the DT for all customers is that customers interested in Recipes have a high-click rate, this explains the first group of customers who are interested in receiving newsletters related to recipes identified by clustering.

4.3 Recommendations

The objective of this research is to identify homogenous groups of customers which are good candidates to experiment with in order to increase the overall response rate. One could of course always experiment with those customers which currently are not/low clicking any emails. However, this would not lead to homogenous groups and the heterogeneity present could obscure the effects of the experiments. For this reason, we suggest the methodology outlined above. The fact that all three clusters reveal a different decision tree indicates the benefit of the clustering step. Fig. 6 shows the decision tree model for the whole data, i.e. without a prior clustering step. It's clear that it identifies less potential experimenting segments.

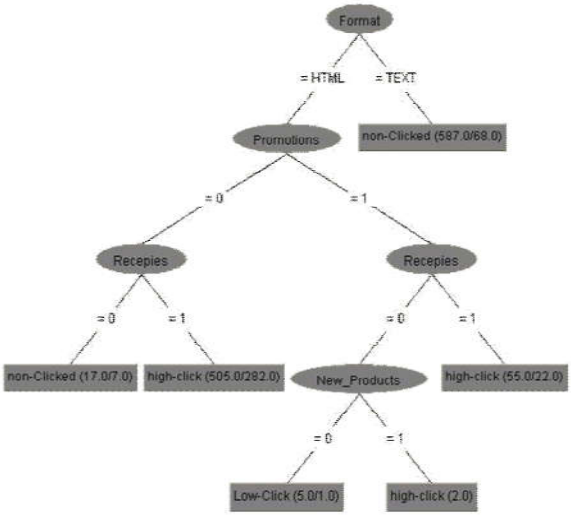

Fig. 6: Decision Tree for All Customers

Based on the results of our analysis, we can formulate the following recommendations:

- Convince customers of cluster 1 (i.e. only interested in recipes) to change their choice of receiving TEXT emails to receive HTML emails. This proves the fact that TEXT format emails are not motivating because it does contain images or videos.
- Filter out customers which receive HTML newsletters and who are likely to be interested in all categories (cluster 2) but did not choose to receive information about. Try to change the layout or other aspects of the newsletters for this group of customers
- Change the sending time of newsletter for customers of cluster 3 receiving HTML emails from afternoon to late evening.
- Change the sending time of newsletter sent in the evening for customers of cluster 3 which are not interested in recipes to the late evening.

CONCLUSIONS AND FUTURE WORK

In this paper, we analysed and examined customers receiving weekly newsletters as a part of an email marketing campaigns, the data studied was from a leading email marketing solution provider in Belgium, the aim of our study is to identify customers who have non/low-click behaviour to allow companies to experiment with those customers. Our methodology of analysis has been performed in two steps, first by identifying homogenous groups of customers according to interests, and step two by applying decision tree analysis as a segmentation technique for each cluster using the click rate categorized as the class variable.

After identifying target customers to be experimented for increasing the response rate, we recommended some actions to be taken to those customers. The future work will be to set up experiments for the identified candidate groups and to evaluate the effect of these experiments on the overall click rate.

REFERENCES

[1] Ansari Asim, Mela Carl F. E-customization. J Mark Res 2003;40(2):131–46.
[2] Arabie, P., Hubert, L.J., 1994. Cluster Analysis in Marketing Research. In Advanced Methods of Marketing Research, R.P. Bagozzi ed. Oxford: Blackwell, 160-189.
[3] Everitt B., Landau, S., Leese, M., 2001. Cluster Analysis. Arnold. 22.
[4] Fayyad, U.M, Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Eds.), Advances in Knowledge Discovery and Data Mining. AAA1 Press/The MIT Press, pp 1-34.
[5] Fraley, C., Raftery, A.E., 2002. Model-based clustering, Discriminant Analysis, and Density Estimation. Journal of the American Statistical Association 97 (458), 611-631.
[6] Godin, S. (1999). Permission marketing: Turning strangers into friends and friends into customers. New York: Simon Schuster.
[7] Godin, Seth (2000), "Permission Marketing," Credit Union Executive, 41 (January), 42.
[8] Green, P.E., 2004. Practice makes perfect. Marketing Research 16 (2), 8-14.

[9] Han, Y., Hu, C., Bal, B., & Jang, S. (2005). Marketing hospitality and tourism education on the Internet: An analysis of e-relationship marketing features. Journal of Hospitality and Tourism Education, 17(1), 11–22.

[10] Hair J. F. Jr., Anderson, R.E., Tatham, R.I., Black, W.C., 1998, Multivariate Data Analysis, Prentice Hall.

[11] Jiang, D., Tang, C., Zhang, A., 2004. Cluster analysis for gene expression data: a survey. IEEE Transactions on Knowledge and Data Engineering 16 (11), 1370-1386.

[12] J. Ross Quinlan , 1992, C4.5 programs for machine learning, Morgan Kaufmann Publishers, ISBN: 1-55860-238-0.

[13] Kent, R., & Brandal, H. (2003). Improving email response in a permission marketing context. International Journal of Market Research, 45(4), 489–503.

[14] Moustaki, I., Papageorgiou, I., 2005. Latent class models for mixed variables with applications in Archaeometry. Computational Statistics and Data Analysis 48 (3), 659-675.

[15] Rust RT, Espinoza F. How technology advances influence business research and marketing strategy. J Bus Res 2006; 59: 1072–8.

[16] Walrave, M. (2004). Met uw Permissie: E-mailmarketing en de opt-in regel. Antwerpen: De Boeck nv.

[17] Vermunt, J.K., Magidson, J., 2005. Technical Guide for Latent GOLD 4.0: Basic and Advanced. Belmont Massachusetts: Statistical Innovations Inc.

★★★