FUZZY MODELING OF CLIENT PREFERENCE IN DATA-RICH MARKETING ENVIRONMENTS

MAGNE SETNES AND UZAY KAYMAK

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2000-49-LIS
Publication	August 2000
Number of pages	10
Email address corresponding author	Kaymak@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam PoBox 1738 3000 DR Rotterdam, The Netherlands Phone: # 31-(0) 10-408 1182 Fax: # 31-(0) 10-408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	Advances in computational methods have led, in the world of financial services, to huge databases of client and market information. In the past decade, various computational intelligence (CI) techniques have been applied in mining this data for obtaining knowledge and in-depth information about the clients and the markets. This paper discusses the application of fuzzy clustering in target selection from large databases for direct marketing (DM) purposes. Actual data from the campaigns of a large financial services provider are used as a test case. The results obtained with the fuzzy clustering approach are compared with those resulting from the current practice of using statistical tools for target selection.	
Library of Congress Classification (LCC)	5001-6182	Business
	5201-5982	Business Science
	HF 5415.126	Direct Marketing
	Q 276-280	Mathematical Statistics
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 11	Production Management
	R 4	Transportation Systems
	M 31	Marketing
	C 8	Data collection and data estimation methodology
European Business Schools Library Group (EBSLG)	85 A	Business General
	260 K	Logistics
	240 B	Information Systems Management
	290 R	Direct Marketing
	250 B	Fuzzy theory
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.34	Logistiek management
	85.20	Bestuurlijke informatie, informatieverzorging
	85.40	Marketing
	31.73	Mathematische Statistiek
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Bedrijfsprocessen, logistiek, management informatiesystemen	
	Direct Marketing, Data mining, Fuzzy theorie, Financiele Instellingen	
Free keywords	Direct Marketing, client segmentation, fuzzy systems, fuzzy clustering	

Fuzzy Modeling of Client Preference in Data-Rich Marketing Environments

Magne Setnes and Uzay Kaymak

Abstract—**Advances in computational methods have led, in the world of financial services, to huge databases of client and market information. In the past decade, various computational intelligence (CI) techniques have been applied in mining this data for obtaining knowledge and in-depth information about the clients and the markets. This paper discusses the application of fuzzy clustering in target selection from large databases for direct marketing (DM) purposes. Actual data from the campaigns of a large financial services provider are used as a test case. The results obtained with the fuzzy clustering approach are compared with those resulting from the current practice of using statistical tools for target selection.**

Keywords—**Direct marketing, client segmentation, fuzzy systems, fuzzy clustering.**

I. Introduction

Large amounts of data are nowadays available to companies for targeted interaction with their customers. Information can be extracted from this data, which can be used for tailoring the products and the service a company provides to the specific requirements of individual customers. Direct marketing companies exploit the information that can be obtained in data rich marketing environments for making selective offers to their customers. Direct marketing is an effective means of interacting with the customer in the field of financial services, and it is being adopted by a growing number of financial services companies as their main strategy for interacting with their customers.

The field of financial services is largely practiced by personnel who rely on proven methods and experience. However, there is also the need to better understand and model the financial processes which are becoming increasingly more complex, which leads to the increased use of computational intelligence (CI) to process the mass of data produced in this field. Interest in the modeling of complex processes has increased in recent years. For engineering problems, CI-based approaches such as fuzzy systems are being applied to describe the qualitative working of complex processes. In classical financial engineering and econometry, however, the trend is still to search for complex, detailed descriptions of systems, despite the fact that the uncertainty and complexity of the studied systems surpass those studied in engineering. In other words, financial engineering has a prospect to benefit from improvements in efficiency, understanding and probably also model accuracy by considering the use of recently emerging CI techniques, such as fuzzy set based techniques.

Considerable and often unwarranted idealization is necessary when modeling business related processes by traditional techniques. As an alternative, fuzzy systems can be used to model many of these inherently uncertain processes. Unlike other CI techniques like neural networks, where the learned knowledge is represented in a network of weights, the possible linguistic interpretation of the rules in a fuzzy system makes this learned knowledge accessible to domain experts [1]. This allows for validation, correction and extension of the knowledge by domain experts, and increases the general acceptance of a fuzzy system by them, which is a prerequisite for the successful introduction of new methodology [2]. The model transparency combined with good function approximation capabilities are the main justifications for applying fuzzy systems in finance and business.

One of the important problems in direct marketing is the selection of customers from a database who will be sent an offer regarding a product. This problem is known as “target selection” in direct marketing. In this report, we propose a structured data mining approach to the target selection problem using fuzzy clustering and fuzzy decision making. The model obtained from the approach identifies the customers who are more likely to respond to a specific product offer, so that they can be targeted preferentially.

The outline of the report is as follows. Section II describes the target selection problem in direct marketing, and discusses the current target selection practice. In Section III, a structured approach to data mining for target selection is proposed. Fuzzy decision making is used in a preprocessing step to deal with missing values in a data set, and for initial reduction of dimensionality. Then fuzzy clustering is used to determine the target selection capacity of each variable in the data base and a hierarchical fuzzy model structure that also incorporates feature selection is obtained. To help overcome the size of the clustering problem, an extended fuzzy clustering algorithm is applied where the cluster prototypes have been extended from single points to variable size sets. The algorithm also automatically determines a suitable partition for each model variable without the need for user intervention or cluster validity measures. In Section IV we use the proposed approach to derive a transparent and compact client model for the purpose of target selection. Actual banking data is considered, and the approach is demonstrated on a marketing campaign of a financial product using a client database with 170 features. The results are compared with those obtained with a method from the current practice of using statistical tools, and it is shown that the fuzzy approach gives about 20% improvement in target recognition. The impact of the results and various aspects to look at in applications are discussed in Section V, before Section VI gives the conclusions.

This work was supported in part by The Research Council of Norway.

M. Setnes is with Heineken Technical Services, Research and Development, Burg. Smeetsweg 1, 2382 PH, Zoeterwoude, the Netherlands, e-mail: magne@ieee.org

U. Kaymak is with Erasmus University Rotterdam, Faculty of Economics, Department of Computer Science, P.O. Box 1738, H9–11, 3000 DR Rotterdam, the Netherlands, e-mail: u.kaymak@ieee.org

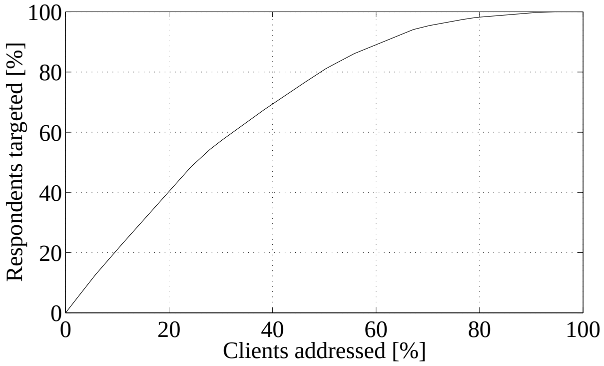
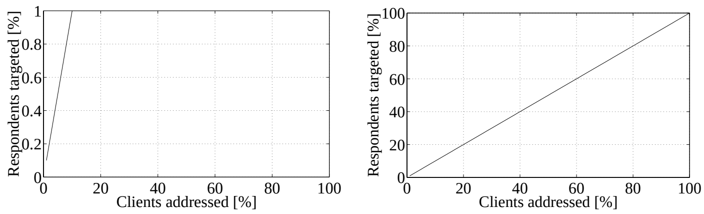


Fig. 1. An example gain-chart. It is seen that the model targets 40% of the respondents by addressing 20% of the clients.



(a) Optimal target selection

(b) Random mailing

Fig. 2. Illustrations of a gain chart corresponding to (a) an optimal target selection model in a data set that contains 10% respondents, and (b) a gain chart resulting from a random approach (no model).

II. Target Selection in Direct Marketing

A great deal of consumer preference information as well as behavior and demographic data are available to many companies. Besides the well known utilization of such data by retailer and producers, many leading credit, bank and insurance companies are also mining this data to customize their products according to client needs and to better target their various marketing campaigns in order to reduce costs. Not only the costs of addressing all the clients in a campaign are of importance, but equally important is the fact that client annoyance due to undesirable mail can lead to loss of market. Especially in the case of internet-based business and email campaigns, not the costs, but client annoyance is the main concern.

In direct marketing, models (profiles) are generated to select potential customers (from the client database) for a given product by analyzing data from similar campaigns, or by organizing test mail campaigns. The models are evaluated by means of gain-charts like the ones shown in Fig. 1. A gain-chart shows the gain made by using a derived model for target selection over random selection of targets. The x-axis shows the ratio (in percentage) of the size of mailed group of clients to the size of the total client base. The y-axis gives the percentage of responders within the group selected for mailing compared to the total number of responders in the client base. Thus, the gain chart shows how many of the positive respondents in the analyzed (test) campaign would have been targeted by the model for any given campaign size. The model in Fig. 1 shows that with a campaign directed to 20 % of the clients, 40% of the respondents are approached, as opposed to only 20 % in a randomly targeted mailing of the same size. A steeply increasing gain-chart is desirable, i.e. the model should target as many respondents as possible while addressing as few clients as possible. For instance, if in a given data set 10% of the clients are respondents, an optimal model will target all these by approaching 10% of the clients, i.e. all clients addressed are respondents. This is the ideal case, where addressing more clients will not benefit the efficiency of the campaign. This is shown in Fig. 2a. Conversely, a straight line from lower left to upper right corresponds to a random mailing, i.e. no model is used. This is illustrated in Fig. 2b. The goal of target selection is to determine a target selection model such that the gain chart corresponding to the model shifts from the one in Fig. 2b towards the one in Fig. 2a.

The analysis of a direct mail campaign entails two stages. First, *feature selection* must determine the variables that are relevant for the specific target selection problem. Second, the rules for selecting the customers should be determined, given the relevant features. At present, statistical tools like CHAID [3] are used to search for features that discriminate between respondents and non respondents, and to build decision tree models for target selection. CHAID (CHi-Square Automatic Interaction Detection) uses the χ^2–tests [4] to identify the targets in a client data base.

A point of concern for target selection methods is the over-training of the model. Over-training implies that the model fits the data used for training very well, but may give worse results when applied to unseen data. Hence, the generalizing power of the resulting model is small. In statistical approaches, the over-training may be due to the sensitivity of the model to the classification of features into crisp sets. Moreover, the number of respondents in a training data set is typically very small compared to the total number customers, which biases the results and the applicability of the assumed distributions. Due to the rather non-transparent nature of many statistical models, over-training can not be detected by expert inspection. The typical solution is to use an evaluation data set to ensure the generalizing capabilities of the model. However, especially in case of test mail campaigns, this introduces extra costs as more mailings must be made in order to gather reasonably sized separate training and evaluation data sets. Also, the time constraint does not always allow for evaluation. Thus, well generalizing models that are not easily over-trained, and that allow for some kind of expert inspection are needed. These are both properties of fuzzy set models.

Encouraged by results obtained in other fields, fuzzy clustering is chosen as a tool for the modeling task in target selection. In addition to good discrimination properties, fuzzy clustering also offers the possibility of extracting linguistic rules from the fuzzy clusters by letting each cluster correspond to a decision rule, and describe it by means of fuzzy sets. This gives transparent models which enable evaluation and tuning by domain experts, increasing the acceptance of new methodologies.

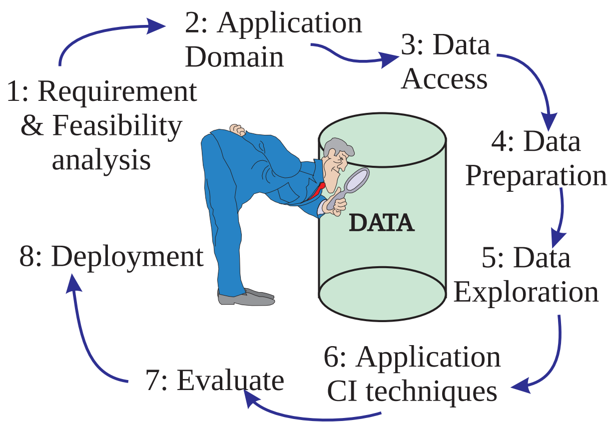


Fig. 3. A general cycle of a data mining project using CI techniques.

III. Fuzzy Data Mining and Modeling for Target Selection

In this section we propose a structured approach to data mining for target selection. The general cycle of a data mining project is depicted in Fig. 3. In the following, we will focus attention on steps 3 to 7. Although not discussed here, it is important to remember that for the successful introduction of any new CI technique in business, the requirement and feasibility analysis is crucial to the selection of the application domain, as well as the proper guidance and post evaluation in the actual deployment.

A. Data access

The data used for constructing direct marketing models are collected from a client database. In this paper, a data set regarding a savings product is considered from a direct marketing campaign of a large Dutch financial services provider. For a similar product, the interest is to obtain a customer preference model that can be used for target selection. The campaign used for training the model contains 16525 clients (data points) described by 170 features (variables) and a response indication. The features consist of 61 binary, 54 continuous and 55 categorical variables. Binary variables typically indicate the possession of a product, such as the possession of a credit card or an investment account. Continuous variables are typically monetary values associated with accounts and products, while categorical variables denote a small quantization (classification) of a continuous variable, such as age (divided into eight classes), and various classifications of the clients based on their possessions and their relation with the bank. In the data set studied, 1504 clients (9.1%) have responded to the campaign.

B. Data preparation

Typical for the application domain, the data set contains a lot of missing values. Only the binary variables that concern the possession of the bank's products are complete. If a customer does not possess a product, the corresponding monetary value (balance) is not filled in. Hence, the corresponding continuous variable will have a missing value in it. By simply filling in "0" as balance for non-possessed products, a lot of missing values can be accounted for.

In order to get a complete data set, practitioners often replace missing values with a number that corresponds to a classification "don't-know". While this can be acceptable for some statistical approaches, other methods, like clustering in metric space, requires a complete data set. Apart from a model-based completion of the data, this can be obtained in two ways: features with missing values can be removed from the data set, or, similarly, data points corresponding to clients with uncompleted records can be removed. These two approaches (removal of features and removal of records) are contradictory, and hence a trade-off between the conservation of the features and the conservation of records is needed. On one hand, one prefers to retain as many customer records as possible, so that sufficient number of examples are available for all interesting customer groups. On the other hand, one wants to consider as many features as possible during modeling in order to capture all relevant relations in the data set. Assume now that a threshold $\eta \in [0,1]$ is defined such that all features in the data set with the proportion of missing values larger than η are removed. Afterwards, the records still containing missing values are removed from the data set. The question is now determining the optimal value for η.

The optimal threshold η^* can be determined using a fuzzy decision making approach. A fuzzy set "most records retained" can be defined on the universe of allowed percentage of missing values in a feature. Typically, small percentage of allowed missing values will have a high membership in this set. Additionally, a fuzzy set "most features allowed" can be defined also on the universe of allowed percentage of missing values in a feature. Large percentage of allowed missing values will typically have a high membership in this set. Given these two fuzzy sets (fuzzy goals), the optimal threshold is found using Bellman and Zadeh's fuzzy decision making model [5], where the maximizing decision is taken as the optimal threshold that satisfies both goals. In our case, the fuzzy sets for "most records retained" and "most features allowed" are derived from the data by considering the number of missing values in each feature and the resulting decrease in the number of features and records. Furthermore, both criteria are assumed to be equally important. The optimal threshold η^* is now determined as the point where the intersection of the two fuzzy goals has its maximum membership. The fuzzy goals derived from the direct marketing data considered in this report are shown in Fig. 4. The maximizing decision in this case is to allow for 5% missing information in a feature.

Having determined the allowed percentage of missing information in a feature, all features with a higher percentage of missing values are removed. Thereafter, client records still having missing values in this reduced feature space are removed from the data set. The result is a reduced but complete data set. An illustration is given in Example III.1.

Example III.1: Consider the initial 6×4 pattern matrix shown in Fig. 5. Removing all client records $Ck, k = 1, 2, \ldots, 6$ with missing features X would leave only one client in the data set, while removing all features $Fj, j = 1, 2, 3, 4$ with missing values would leave only one feature in the data set. The goal "most features allowed" is obtained by determining the number of features retained in the data set, provided a given number of missing values is allowed in a feature. The goal "most records re-

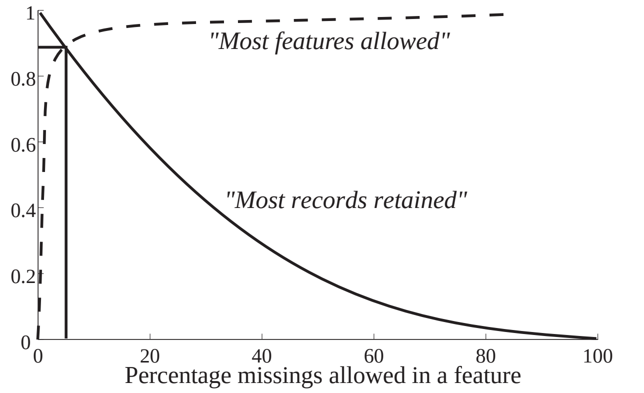


Fig. 4. When 5 % missing information is allowed in a variable, 148 features and 15486 client records are retained in the data set.

tained" is obtained by determining the number of records that are retained after deleting the non-qualifying features and removing the records that still contain missing values. These fuzzy goals derived from the data are shown in the graph in Fig. 5. The optimal decision is found from the intersection of the fuzzy sets. In this example, allowing 2 missing values in a feature leads to the optimal decision, retaining 2 features (50%) and 4 client records (66%).

In the direct marketing data considered in this report, η^* is equal to 5% as illustrated in Fig. 4. When the data set is completed by reduction as described above, the result is that 22 features are removed from the data set and 1039 client records are deleted.

C. Data exploration

After the data preparation, the data set is ready to be explored for structures that can be used by the model. This is done by means of fuzzy clustering [6], and for this reason, it is necessary that the features span a metric space. In order to obtain this, the binary and categorical features must be removed from the data set. Binary variables indicate product possession. After having completed the balance information in the previous step, all information about product possessions are present in the balance information in the corresponding continuous variables. Removing the binary possession variable corresponds to a projection of the data onto the balance variable. No information is lost in this step. The categorical information is mostly distilled from combinations of possession and balance information, and hence this feature reduction is also justified.

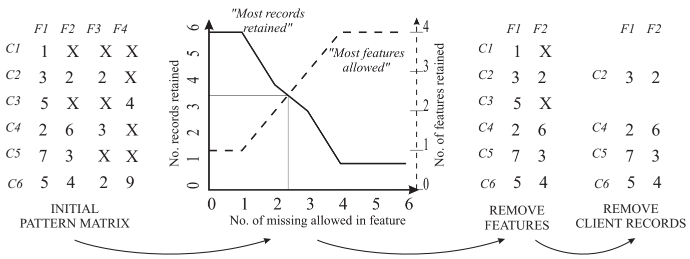


Fig. 5. Decision making approach for completing the data set by reduction. First, derive the fuzzy goals and determine the maximizing decision. Second, remove the features with too many missing values. Third, delete the remaining incomplete client records.

After having obtained a complete data set by reduction, and having removed the binary and categorical features, the dimensionality of the data set is reduced to 49 features. However, this is still a too high dimensionality. A model with this dimensionality is not very transparent, and, furthermore, clustering in such a space is tedious. Thus, further reductions are necessary.

For the purpose of feature selection, a number of methods are studied for comparison purposes. Linear discriminant functions and modified principal component analysis [7] have been used to select variables for the clustering based modeling method. Results obtained with these methods, however, are poor. In general, linear feature selection methods are less suitable due to the non-linearities and inconsistency in the direct marketing data. A third method that has been tested is the so-called Y-association which is also used by CHAID [3]. This method is derived from the Chi-square test and makes use of an assumed statistical distribution which can hardly be said to be present in the data. Nevertheless, the results obtained are better than those obtained when using the two previous techniques.

A novel approach is to do feature selection using gain-charts resulting from individual models made for each variable. The idea is derived from the way in which the domain experts analyze direct marketing models, described in Section II. Compared to other feature selection methods, this provided the best results and is applied in the modeling described in the next section. The modeling itself is a step-wise refinement of the gain-chart produced by the model, starting with the most favorable feature, and including a new feature (input variable) for each step in a decision tree like manner. This allows for the feature selection to be inherent in the modeling. A feature is selected and added to the model at each level in the tree, and the number of features used depends on the depth of the resulting tree. The depth can be determined by user, or the modeling can be set to terminate at a non-predetermined depth depending on its targeting properties. This feature selection is described below.

D. Modeling

Due to the high dimension and sparse population of the space spanned by the variables, product space clustering is not feasible. The nature of the data is such that much information is conveyed by considering single variables in an isolated manner. The problem of modeling is thus to distill and combine this information in a proper manner. To achieve this, we propose an incremental modeling approach based on step-wise refinements. The modeling is done for each variable using fuzzy clustering in one dimension and gain analysis, leading to additive improvement of the gain charts per feature and of the total model as models of individual features are added to the total model in a decision tree like manner.

D.1 Ranking clients per feature

Let the training data set be $\{\mathbf{x}_k, y_k\}, k = 1, 2, \ldots, N$, where $\mathbf{x}_k = [x_{k1}, x_{k2}, \ldots, x_{kn}]$ is the n-dimensional feature vector describing the k'th client (data object) and $y_k \in \{0, 1\}$ is the corresponding response label. The n features are studied separately to see how well they discriminate between the two groups in the data (respondents and non-respondents). The samples x_{jk} of each feature $X_j, j = 1, 2, \ldots n$ are partitioned into M_j clusters by means of fuzzy clustering, and the fuzzy Respondent Density (RD) of the resulting clusters is computed. The fuzzy respondent density is given by

$$\mathrm{RD}_{ij} = \frac{\sum_{k=1}^{N} \mu_{ij}(x_{kj}) y_k}{\sum_{k=1}^{N} \mu_{ij}(x_{kj})}, 1 \le i \le M_j, j = 1, 2, \ldots, n, \tag{1}$$

where M_j is the number of clusters identified for feature X_j and $\mu_{ij}(x_{kj})$ is the membership of the k'th client in the i'th cluster of feature j. Thus, RD_{ij} is the ratio of the total membership of the positive responders in the cluster to the total membership of all clients in the cluster.

The RD_{ij} values are now used to compute a score (SC_{jk}) for each client record $\mathbf{x}_k$ according to the feature X_j:

$$\mathrm{SC}_{jk} = \sum_{i=1}^{M_j} \mu_{ij}(x_{kj}) \mathrm{RD}_{ij}, 1 \le k \le N. \tag{2}$$

The higher the score SC_{jk}, the more plausible that the k'th client is a respondent, according to feature j. Thus (2) represents a model that ranks the clients according to the j'th feature. In this way, a model for each feature X_j is obtained, where the M_j cluster prototypes resulting from the partitioning of the data and the corresponding RD values calculated by (1) are the parameters constituting the model.

D.2 Iterative modeling

The model obtained for each feature can be evaluated by gain-chart analysis. To make a gain chart for a feature X_j, first all clients are sorted according to decreasing SC_{jk} values ("best client first"). A gain chart is now obtained by growing a target set from 1 to N clients, whereby clients are added in the sorted order k^*, and the cumulative sum of the corresponding response y_{k^*} is recorded. The gain chart value is then given by

$$g_j\left(\frac{k*}{N}\right) = \frac{\sum_{i=1}^{k^*} y_i}{\sum_{i=1}^{N} y_i}, \; k^* = 1, 2 \ldots N. \tag{3}$$

The gain chart for a feature X_j is assigned a score SX_j which reflects its targeting efficiency by weighting it for any k^* with the fraction of clients not yet added to the target set:

$$SX_j = \sum_{k^*=1}^{N} \left(1 - \frac{k^*}{N}\right) g_j\left(\frac{k^*}{N}\right). \tag{4}$$

The weighted scoring favors gain charts with a steep rise in the beginning. This gain chart behavior is preferred in target selection, since the goal is to target as many respondents as possible with as few mailings as possible.

Using gain chart analysis, the modeling proceeds in an iterative manner that incorporates feature selection. For $l = 1, 2, \ldots, D$ iterations, where D is the depth of the resulting decision tree (and also the number of features eventually used in the final model), the gain-charts produced by each of the individual features in the total feature set are compared according to their score SX_j. At each iteration l of the modeling process, the feature having the most favorable gain-chart is selected, i.e.

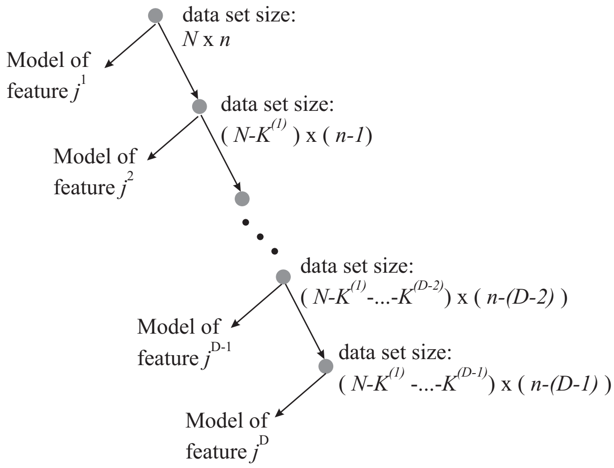


Fig. 6. In each step, a target model is produced from all features in the data set. The parameters corresponding to the feature model with the best gain-chart are recorded. Then the size and dimension of the data is reduced, and the modeling process repeats on the reduced data set.

$j^l = \mathrm{argmax}_j SX_j$. The M_{j^l} cluster centers for the selected feature and the corresponding RD_{ij^l} values are recorded as part of the final model. Now, the client data set is partitioned into n' groups $K^{j'}$, $j' = 1, \ldots, n'$ such that

$$K^{j'} = \{k | SC_{j'k} = \max_j SC_{jk}\}. \tag{5}$$

Each group $K^{j'}$ now contains clients that are described best by the feature j'. Next, the $K^{(l)}$ clients who are described best by j^l, i.e. by the feature with the most favorable gain chart, are removed from the training data set, i.e.

$$K^{(l)} = |K^{j^l}|, \tag{6}$$

where $|\cdot|$ denotes the cardinality of the set. The dimensionality of the data set is then reduced by removing the feature j^l. The procedure is then repeated at the next iteration $(l+1)$ on the reduced data set. The process is illustrated in Fig. 6.

By iteratively reducing the data set, the most important structures in the data that can be described by a single feature are sought. At step l, one feature may partition nicely into a few relatively distinctive classes, of which one might have a rather high respondent density value. When the clients having a higher score on the model based on this feature are removed from the data set, other structures in the data can reveal themselves in the partitioning of one of the remaining features. This makes it possible to capture some of the structures in the data for which one normally would have to consider the interactions between features (see Example III.2). The number of iterations is equal to the depth of the resulting decision tree, and the user selects this by indicating the number of features desired in the resulting model. It is not necessary to pre-specify the depth D. The modeling process can also be terminated based on criteria such as "there are no respondents left in the remaining training set". This would select all the features necessary to fully describe all responders in the data set.

Example III.2: In Fig. 7, a problem with two features, X_1 and X_2 is considered with the data shown in the product space

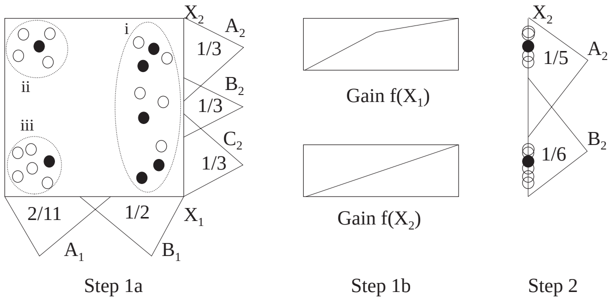


Fig. 7. By iteratively reducing the data set and repeating the modeling, it is possible to describe structures in the data that would normally require product space modeling. This is especially suitable when the classes of data are not well separated like in direct marketing.

of the two features. In step 1a, both features are clustered individually, giving two clusters for X_1 (A_1 and B_1) and three for X_2 (A_2, B_2 and B_3). In step 1b, the evaluation of the gain charts of the two resulting models $f(X_1)$ and $f(X_2)$, based on X_1 and X_2 respectively, reveals that X_1 is the best feature. The target model $f(X_1)$ is recorded, i.e. the cluster prototypes and the corresponding RD values, and modeling is repeated on a reduced data set in step 2. In the reduced data set the feature X_1 has been removed, and the data objects that scored better on the model $f(X_1)$ than on $f(X_2)$ have also been removed. That is, the data marked group i in step 1a are removed. This data scored 1/2 on model $f(X_1)$ and $1/3$ on model $f(X_2)$. The data marked group ii and iii scored respectively 2/11 and 1/3 on the models $f(X_1)$ and $f(X_2)$ in step 1. Data groups ii and iii remain in the reduced data set, and two clusters are found for X_2 in step 2.

D.3 The final target-selection model

After having iteratively created a target model for the D most attractive features as described above, these individual models are combined in a final model. The ranking and the selection of the clients is done by averaging the results from each feature model. The target selection score for a client described by record $\mathbf{x}_k$ is given by

$$SC_k = \frac{1}{D}\sum_{j^l=1}^{D} SC_{kj^l} \ . \tag{7}$$

All the clients in the data set are now ranked according to their SC_k value. The ones with the highest SC value are targeted. The actual decision on where to put the cut-off threshold is taken by the campaign organizer, and it depends on a number of factors that make up the specific campaign objective, such as company strategy, business plan, the profitability of product being marketed, costs and the efficiency of the information carrier. This aspect is not considered further in this report.

D.4 The clustering

In order to obtain a model for a feature, fuzzy clustering is used to partition the data into a number of partly overlapping clusters based on their distance measured along the feature. For the considered problem, it is not only important to determine the number of clusters in the data, but the algorithm should also be able to find clusters containing a highly varying number of data objects. To deal with these issues, we apply an extended version of the fuzzy c-means (FCM) [6] algorithm that incorporates cluster merging and variable sized cluster prototypes [8].

An important issue when applying clustering is the determination of the relevant number of clusters in the data. A method that is applied often in the literature is the repetitive clustering with various number of clusters and partition assessment with cluster validity measures [9]. This approach is not feasible for the application considered due to the large number of features and the repeated modeling in each iteration. Moreover, the nature of the data is such that the validity measures typically give no insight into the suitable partitioning as there is usually no distinctive local minima of the validity function. To overcome this problem, the extended FCM (E-FCM) algorithm applies a cluster merging approach based on cluster similarity [8]. The algorithm is initialized with a larger number of clusters than what is assumed to be needed. As the clustering progresses, it automatically determines a suitable partition by iteratively merging pairs of similar clusters. The similarity assessment is based on the fuzzy set inclusion measure, and the pairwise similarity between two clusters A and B is determined as

$$S(A,B) = \max\left(\frac{|A\cap B|}{|A|}, \frac{|A\cap B|}{|B|}\right), \tag{8}$$

where $|\cdot|$ denotes the cardinality of a fuzzy set, and $\cap$ represents the intersection, modeled by the $\min$ operator [10]. The similarity $S(A,B)$ is thus given by the maximum of the inclusion of A in B and the inclusion of B in A. $S(A,B)$ is a set-theoretic similarity measure that is not influenced by the numerical scale of the clustered data [11].

This approach was first proposed in [8], where a similarity threshold for cluster merging $\lambda \in [0,1]$ had to be set by the user. This threshold can sometimes be rather difficult to determine, as the cluster similarity is strongly influenced by the number of clusters present in the partition at any time due to the membership constraint applied in the clustering algorithm[6]. In the application of the algorithm to various problems, it has been found that an adaptive threshold, depending on the number of clusters, gives satisfactory results [12]. This approach is also taken here, and the similarity threshold for cluster merging is $\lambda = 1/M$, where M is the number of clusters in the partition at any time.

Another feature of the marketing data is that the ratio of responders to non-responders can be very low at any step of the iterative modeling process. Since the responders might be characterized by feature values for which there are only few representative values, the clustering algorithm must be able to detect clusters of relatively small size (volume) as well as bigger ones in the same data set. The E-FCM algorithm achieves this by applying volume prototypes as opposed to the single point prototypes in the standard FCM algorithm. Moreover, the sizes of the prototypes depend on the fuzzy covariance of the data in its proximity. The effect is that clusters in regions with little data will contain these data objects with membership one, making the cluster center less likely to move towards other regions of the space that contains more data. When many clusters are present in the same part of the data space, the volume prototypes will spread, making it more likely that a cluster prototype will move to another part of the data space that is populated by

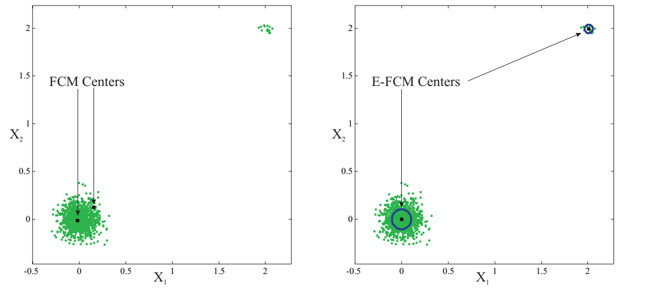


Fig. 8. Clustering data with two groups represented by 1000 and 15 points respectively. Results obtained for two clusters with FCM (left) and with E-FCM (right).

fewer data objects. The volume prototypes in E-FCM are defined by the data objects $\mathbf{x}_k$ having a distance of zero from the volume prototypes. The distance measure is given by

$$d_{ik} = \max(0, \sqrt{(\mathbf{x}_k - \mathbf{v}_i)^T(\mathbf{x}_k - \mathbf{v}_i)} - r_i)\,, i = 1, 2, \ldots M\,, \tag{9}$$

where $\mathbf{v}_i$ are the cluster centers and r_i is the radius of the volume prototype of cluster i. The radius is computed from the fuzzy co-variance of the data belonging to the cluster:

$$r_i = \sqrt{|P_i|^{1/n}}\,, i = 1, 2, \ldots M, \tag{10}$$

with P_i being the covariance matrix for cluster i. The radius (10) reduces to the standard deviation when clustering in one dimension in the proposed fuzzy target selection approach. All data objects that are within this radius have a distance zero to the prototype, and are as such assigned a membership $\mu_{ij}(\mathbf{x}_k) = 1$.

The working of the E-FCM algorithm is illustrated in Fig. 8. A data set in two dimensions is shown with two groups of uneven size. When the standard FCM is applied with two clusters, the prototypes are both located in the same part of the data space, missing the smaller group in the data. When E-FCM is applied, a "big" cluster will locate itself in the densely populated part of the space, forcing the other cluster to move to the small group of data points. When E-FCM is initialized with an over-determined number of clusters, most clusters will be located in the densely populated region, where they are iteratively merged owing to their similarity. This is illustrated in Fig. 9 which shows an example with four uneven sized groups of data.

Note that the E-FCM algorithm can also deal with categorized and binary data. The volume prototypes then tend to be located on the categories. In this case, our modeling algorithm reduces to a simple scoring algorithm based on the percentage of respondents in each category.

IV. Application Example

We report the results from the application of the proposed modeling procedure to the data obtained from a direct marketing campaign regarding a savings product of a large Dutch financial services company. The initial data set contains 170 features. After the data preparation and reduction described in Section III-B and Section III-C, the data set consist of 15.486 data objects in a 49 dimensional feature space. This data set was split at random into a training data set (10.324 clients) and an evaluation data set (5.162 clients). The iterative modeling process is applied to the training data, using the E-FCM clustering algorithm to partition the data in each feature. The algorithm is initialized with 10 clusters each time.

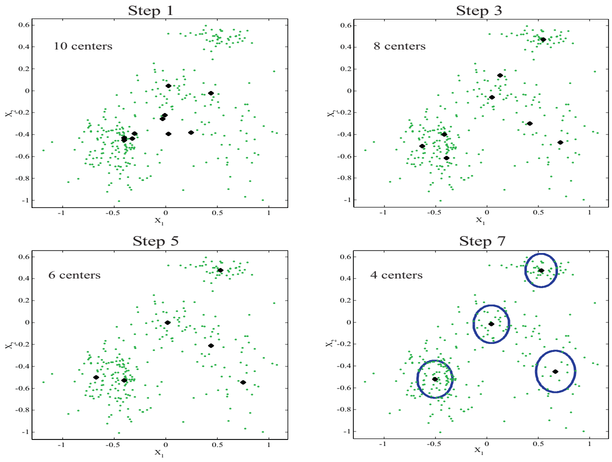


Fig. 9. Excerpts from the progress of the E-FCM clustering algorithm when initialized with 10 random clusters in a data set with four groups.

A. Iterative modeling

The iterative modeling process is set to terminate when there are no more responders left in the iteratively reduced data set. A total of 28 features are selected in decreasing order of importance before the procedure terminates. The number of clusters detected per feature varies between 2 and 9, with a mean of 4.5. The model can now be used with any number $D \leq 28$ of sub-models by considering only the D features first selected in the modeling process. In the following we consider only the first three features. In Fig. 10, the first three steps in the modeling process are illustrated, with the gain charts ordered according to decreasing SX_j score at each step. The three best features in the considered data set are *total savings balance, balance of p-account* and *balance of s-account*. The E-FCM clustering algorithm detects four, three and three clusters for each of these features, respectively.

B. Total model

The recorded cluster prototypes for each selected feature and the corresponding RD_{ij^l} values constitute the parameters of the total model. Clients are scored by the model according to (7) by taking the average of their scores on the individual models constructed for each of the used features *total savings balance, balance of p-account*, and *balance of s-account*:

$$\mathrm{SC}_k = \frac{1}{3}(\mathrm{SC}_k^{savings} + \mathrm{SC}_k^{p-account} + \mathrm{SC}_k^{s-account})\,. \tag{11}$$

Figure 11a shows the gain-charts of the resulting model consisting of the three sub-models whose gain charts are shown in Fig. 10. Both the gain chart for the training data (10.324 clients) and for the evaluation data (5.162 clients) are shown. For comparison, in Fig. 11b, the gain charts resulting from a depth three CHAID model are shown. This model was constructed according to the current practice, using all the 170 features and with

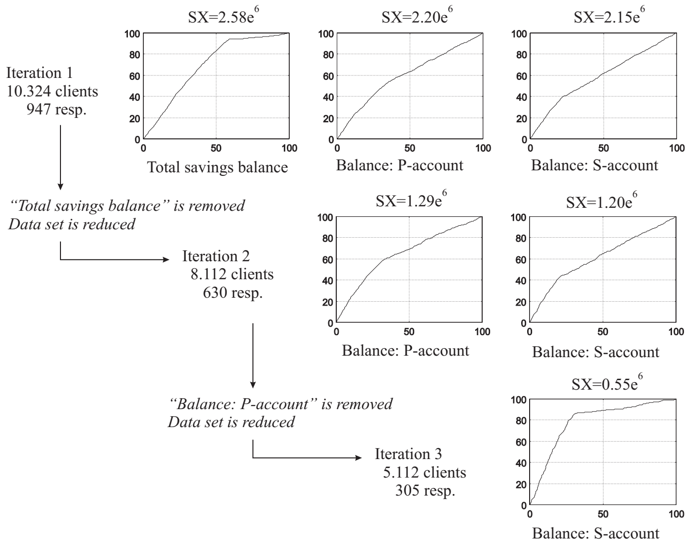
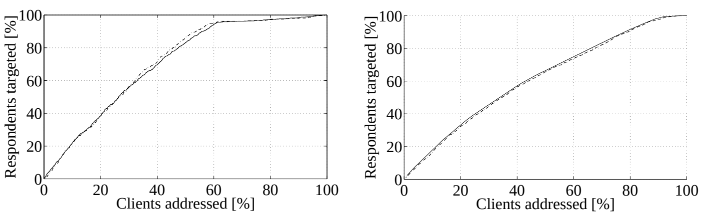


Fig. 10. Gain charts produced during the iterative modeling. In each successive step both the dimensionality and the size of the data set are reduced.

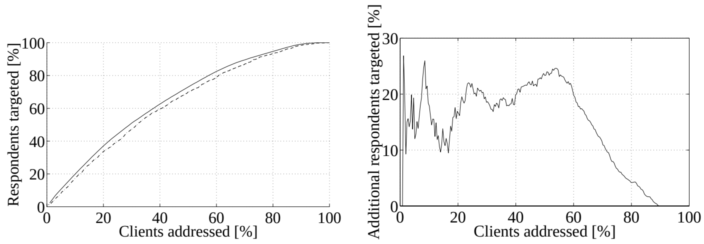
the missing values taken as a separate class. The results obtained with the proposed modeling approach clearly outperform the CHAID model. Moreover, while the fuzzy model of depth three consists of ten simple rules and makes use of three features, the CHAID model has 21 rules and uses 10 inputs.

Even though the depth-3 CHAID model is more complex than the fuzzy model, the time it takes to compute it is a fraction of the construction time of the fuzzy model. The efficient CHAID modeling algorithm can easily handle more complex models. For a fair comparison, we increased the depth till six for the CHAID model. The resulting gain charts are shown in Fig. 12a. For this depth, CHAID selects a total of 25 features and constructs a decision tree with 52 rules. At this depth, the model is starting to show signs of overtrainig as can bee seen from the difference between the training and the evaluation results. Further increase in complexity (depth larger than 6) does not lead to improved performance. The simple fuzzy model still outperforms the depth-6 CHAID model. The improvement is visualized in Fig. 12b which shows the relative improvement in targeting performance (evaluation data set) of the of the fuzzy model compared to the depth-6 CHAID model.



(a) Fuzzy model *(b) CHAID model*

Fig. 11. Gain charts resulting from the models obtained by (a) the iterative fuzzy modeling approach and (b) the CHAID modeling approach with depth 3. Both training data (solid) and evaluation data (dash-dot) are shown.



(a) CHAID model *(b) Improvement with fuzzy model*

Fig. 12. Gain charts resulting from the depth-6 CHAID model (a) and the targeting improvement of the fuzzy model compared to the depth-6 CHAID model, depending on the size of the addressed client set.

C. Rule extraction

To enable easy inspection and validation of the fuzzy model, a rule-based version can be extracted from the parameter-based model obtained through the iterative modeling. This also allows the model to be augmented with other knowledge, and makes the rules accessible for other purposes.

For each feature $X_j, j = 1, \ldots, D$ used in the model, the M_j identified clusters correspond to one rule each. The cluster center determines the one dimensional fuzzy antecedent of the rule, and the associated RD value can be taken to be the rule consequent. The M_j clusters can be replaced with fuzzy sets $A_{ij}, i = 1, \ldots, M_j$. One way of doing this is to associate each cluster with a trapezoid fuzzy set where the volume cluster prototype corresponds to the core of the fuzzy set. The left (right) support of the fuzzy set is taken to be the right (left) support of the core of the left (right) neighboring set, or the minimum (maximum) of the domain if the set is the leftmost (rightmost) in the partition. In this way a so-called normal fuzzy partition [13] is obtained. The rule-based model contains a total of $\sum_{j=1}^{D} M_j$ rules. For the studied model, a rule base with 10 rules is obtained with rules of the following form:

$$R_i : \textbf{If} \text{ feature } X_j \text{ is } A_{ij} \textbf{ then} \text{ response } y = c_i \tag{12}$$

where the rules consequents c_i can be taken to be the corresponding RD value, or they can be estimated from the training data.

We apply the Takagi-Sugeno [14] type reasoning, and choose to estimate the rule consequents from the available training data using a local least squares approach [15]. The rule base is listed in Table I, where linguistic terms have been associated both with the rule antecedent and the consequent value. The numerical outputs of each rule calculated by least squares parameter estimation are shown in the table as well, and the fuzzy sets for the antecedents are shown in Fig. 13 together with their linguistic labels.

The CHAID model can also be expressed in a rule-based manner. It builds a decision tree by dividing the various features into classes. The number of leaves (end nodes) on the tree corresponds to the number of rules. As mentioned above, the depth-3 and depth-6 CHAID models use 21 and 52 rules, respectively.

TABLE I

LINGUISTIC RULES EXTRACTED FROM THE MODEL.

R_1	If *Total savings balance* is Low	Then response is Low	(y=0.0463)
R_2	If *Total savings balance* is Moderate	Then response is High	(y=0.1559)
R_3	If *Total savings balance* is High	Then response is High	(y=0.1631)
R_4	If *Total savings balance* is Very High	Then response is High	(y=0.1542)
R_5	If *Balance: P-account* is Low	Then response is Low	(y=0.0697)
R_6	If *Balance: P-account* is Moderate	Then response is Moderate	(y=0.1338)
R_7	If *Balance: P-account* is High	Then response is Very High	(y=0.1839)
R_8	If *Balance: S-account* is Low	Then response is Low	(y=0.0733)
R_9	If *Balance: S-account* is Moderate	Then response is High	(y=0.1608)
$R_1 0$	If *Balance: S-account* is High	Then response is Very High	(y=0.1873)

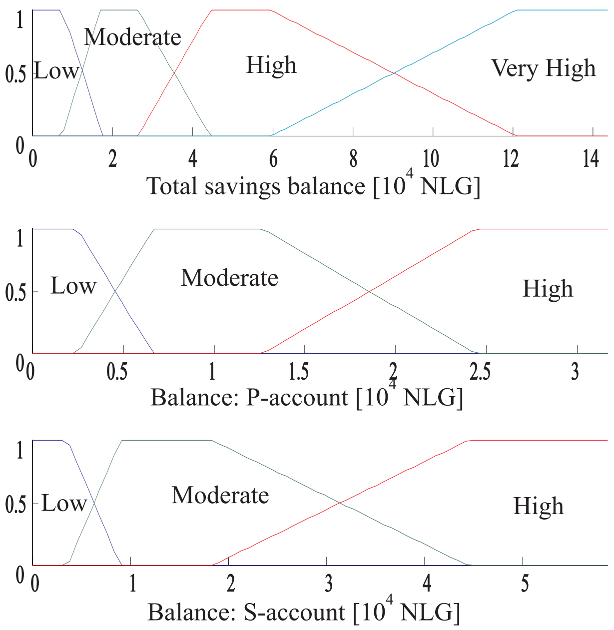


Fig. 13. Fuzzy sets used in the rule-based target selection model.

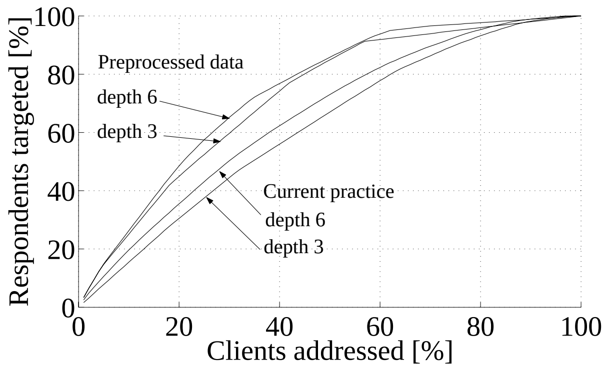


Fig. 14. Results obtained with CHAID models according to current practice, using the total data set with 170 features, and when the data is preprocessed as in the proposed fuzzy modeling approach.

V. DISCUSSION

The success of the proposed fuzzy modeling approach to target selection over the statistically oriented CHAID method from the current practice illustrates the problems sometimes associated with the application of statistical techniques in practice. Especially the assumptions on the distribution of data are typically not valid. This may limit the performance of statistical techniques considerably. A pattern recognition technique such as fuzzy clustering is more adaptive to the specifics of the data as it has few assumptions regarding the distribution of data. The computational load of the cluster based method, however, is orders of magnitude higher than that of the statistical approach. Unlike the CHAID approach where a client belongs to a group with a certain score, in the proposed approach each client gets an individual score. This can help the service provider to tailor products to individual specifics.

Another explanation for the good performance of the proposed fuzzy modeling approach can be found in the data preparation and exploration part. As mentioned in Section III-C, the assumptions for the Y-association test used by CHAID for feature selection might not always be valid. When we apply the CHAID modeling tool to the data preprocessed according to the method in the proposed fuzzy modeling approach, the results improve. Training data results obtained with the current practice CHAID models of depth 3 and 6 are shown in Fig. 14 together with the results obtained when using the preprocessed data set with 49 features only. As it can be seen, the results obtained with the data preprocessed as in the proposed fuzzy modeling approach are superior to the current practice results. Moreover, the number of rules and selected features in the CHAID models based on the preprocessed data set is less that in the CHAID models of the current practice approach. For instance, with the preprocessed data, the depth-3 CHAID model consist of 11 rules and uses five features. The three features used by the fuzzy model are in this case also selected by the CHAID model.

A drawback of all data driven techniques is the need for sufficient data during training. The collection of data for direct marketing campaigns can be costly. However, it is important to collect sufficient and up-to-date data in order to ensure a proper targeting by the resulting model. It is always a question how possible test campaigns should be targeted in order to collect relevant training data. Random selection may provide a good sampling of the total client population, but it is an expensive approach since the number of respondents will typically be small, thus requiring an extensive test campaign. Targeting the test campaign to a pre-selected group, however, may bias the training results and diminish the generalization capabilities of the model. Thus, validation of the assumptions and hypotheses associated with the test campaign is vital, and statistical methods provide valuable tools here that can benefit the analysts.

Acceptance of new methods by analysts is perhaps the single most important factor for successful application of new computational intelligence techniques in financial engineering. Black box approaches often do not find sufficient acceptance as the reason for the outcome from such systems may not be clear to the analyst. Fuzzy systems provide an advantage to many other CI techniques, since the model can be represented in a rule-based form that the analyst can verify against his own knowledge and experience.

VI. CONCLUSION

A fuzzy modeling tool for target selection in direct marketing has been developed. An iterative modeling method based on unsupervized fuzzy clustering is used for identifying target clients from a customer data base for a given product. Using the specifics of data encountered in direct marketing, the proposed method deals with the feature selection in a simple but elegant way, leading to a robust approach. Moreover, simple linguistic If-Then rules can be extracted from the model, allowing easy validation by the analysts. This increases the confidence in the new technique, and is an important issue for the acceptance of the developed approach.

Even though the computational load of the proposed modeling approach by far exceeds that of the current statistical approaches (e.g. CHAID), the targeting improvement, the compactness and transparency of the obtained model make it worth the effort. The target improvement varies from about 15 % till above 20 % for a realistic campaign size.

REFERENCES

[1] M. Setnes, R. Babuška, and H. B. Verbruggen, "Rule-based modeling: Precision and transparency," *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews*, vol. 28, no. 1, pp. 165–169, 1998.

[2] N. W. Bormans, M. Setnes, U. Kaymak, and H. R. van Nauta Lemke, "Application of fuzzy sets to finance and insurance," in *Proceedings of Interfaces'97*, Montpellier, France, May 1997, pp. 189–191.

[3] SPSS Inc., *SPSS CHAID for Windows 6.0*, Prentice-Hall, 1993.

[4] P. G. Hoel and R. J. Jessen, *Basic Statistics for Business and Economics*, John Wiley and Sons, New York, 3 edition, 1982.

[5] R. E. Bellman and Lotfi A. Zadeh, "Decision-making in a fuzzy environment," *Management Science*, vol. 17, no. 4, pp. 141–164, Dec. 1970.

[6] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Functions*, Plenum Press, New York, 1981.

[7] W. J. Krzanowski, "Ranking principal components to reflect group structure," *Journal of Chemometrics*, vol. 6, pp. 97–102, 1992.

[8] M. Setnes and U. Kaymak, "Extended fuzzy c-means with volume prototypes and cluster merging," in *Proceedings EUFIT'98*, Aachen, Germany, Sept. 1998, pp. 1360–1364.

[9] X. L. Xie and G. A. Beni, "Validity measure for fuzzy clustering," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 3, no. 8, pp. 841–846, 1991.

[10] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, New Jersey, 1995.

[11] M. Setnes, R. Babuška, U. Kaymak, and H. R. van Nauta Lemke, "Similarity measures in fuzzy rule base simplification," *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, vol. 28, no. 3, pp. 376–386, 1998.

[12] M. Setnes, *Complexity Reduction in Fuzzy Systems*, Ph.D. Thesis, Delft University of Technology, Dep. of El. Eng., Control Laboratory, 2000, In preparation.

[13] R. Jager, *Fuzzy Logic in Control*, Ph.D. thesis, Delft University of Technology, Dep. of El. Eng., Control Laboratory, 1995.

[14] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modelling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, pp. 116–132, 1985.

[15] J. Yen, L. Wang, and C. W. Gillespie, "Improving the interpretability of TSK fuzzy models by combining global learning and local learning," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 4, pp. 530–537, 1998.

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

Publications in the Report Series Research[*] in Management

Impact of the Employee Communication and Perceived External Prestige on Organizational Identification
Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn
ERS-2000-01-MKT

Critical Complexities, from marginal paradigms to learning networks
Slawomir Magala
ERS-2000-02-ORG

Forecasting Market Shares from Models for Sales
Dennis Fok & Philip Hans Franses
ERS-2000-03-MKT

A Greedy Heuristic for a Three-Level Multi-Period Single-Sourcing Problem
H. Edwin Romeijn & Dolores Romero Morales
ERS-2000-04-LIS

Integer Constraints for Train Series Connections
Rob A. Zuidwijk & Leo G. Kroon
ERS-2000-05-LIS

Competitive Exception Learning Using Fuzzy Frequency Distribution
W-M. van den Bergh & J. van den Berg
ERS-2000-06-LIS

Start-Up Capital: Differences Between Male and Female Entrepreneurs, 'Does Gender Matter?'
Ingrid Verheul & Roy Thurik
ERS-2000-07-STR

The Effect of Relational Constructs on Relationship Performance: Does Duration Matter?
Peter C. Verhoef, Philip Hans Franses & Janny C. Hoekstra
ERS-2000-08-MKT

Marketing Cooperatives and Financial Structure: a Transaction Costs Economics Analysis
George W.J. Hendrikse & Cees P. Veerman
ERS-2000-09-ORG

---

[*] ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Decision Making in Marketing Management
F&A Financial Decision Making and Accounting
STR Strategic Renewal and the Dynamics of Firms, Networks and Industries

A Marketing Co-operative as a System of Attributes: A case study of VTN/The Greenery International BV,
Jos Bijman, George Hendrikse & Cees Veerman
ERS-2000-10-ORG

Evaluating Style Analysis
Frans A. De Roon, Theo E. Nijman & Jenke R. Ter Horst
ERS-2000-11-F&A

From Skews to a Skewed-t: Modelling option-implied returns by a skewed Student-t
Cyriel de Jong & Ronald Huisman
ERS-2000-12-F&A

Marketing Co-operatives: An Incomplete Contracting Perspective
George W.J. Hendrikse & Cees P. Veerman
ERS-2000-13– ORG

Models and Algorithms for Integration of Vehicle and Crew Scheduling
Richard Freling, Dennis Huisman & Albert P.M. Wagelmans
ERS-2000-14-LIS

Ownership Structure in Agrifood Chains: The Marketing Cooperative
George W.J. Hendrikse & W.J.J. (Jos) Bijman
ERS-2000-15-ORG

Managing Knowledge in a Distributed Decision Making Context: The Way Forward for Decision Support Systems
Sajda Qureshi & Vlatka Hlupic
ERS-2000-16-LIS

Organizational Change and Vested Interests
George W.J. Hendrikse
ERS-2000-17-ORG

Strategies, Uncertainty and Performance of Small Business Startups
Marco van Gelderen, Michael Frese & Roy Thurik
ERS-2000-18-STR

Creation of Managerial Capabilities through Managerial Knowledge Integration: a Competence-Based Perspective
Frans A.J. van den Bosch & Raymond van Wijk
ERS-2000-19-STR

Adaptiveness in Virtual Teams: Organisational Challenges and Research Direction
Sajda Qureshi & Doug Vogel
ERS-2000-20-LIS

Currency Hedging for International Stock Portfolios: A General Approach
Frans A. de Roon, Theo E. Nijman & Bas J.M. Werker
ERS-2000-21-F&A

Transition Processes towards Internal Networks: Differential Paces of Change and Effects on Knowledge Flows at Rabobank Group
Raymond A. van Wijk & Frans A.J. van den Bosch
ERS-2000-22-STR

Assessment of Sustainable Development: a Novel Approach using Fuzzy Set Theory
A.M.G. Cornelissen, J. van den Berg, W.J. Koops, M. Grossman & H.M.J. Udo
ERS-2000-23-LIS

Creating the N-Form Corporation as a Managerial Competence
Raymond vanWijk & Frans A.J. van den Bosch
ERS-2000-24-STR

Competition and Market Dynamics on the Russian Deposits Market
Piet-Hein Admiraal & Martin A. Carree
ERS-2000-25-STR

Interest and Hazard Rates of Russian Saving Banks
Martin A. Carree
ERS-2000-26-STR

The Evolution of the Russian Saving Bank Sector during the Transition Era
Martin A. Carree
ERS-2000-27-STR

Is Polder-Type Governance Good for You? Laissez-Faire Intervention, Wage Restraint, And Dutch Steel
Hans Schenk
ERS-2000-28-ORG

Foundations of a Theory of Social Forms
László Pólos, Michael T. Hannan & Glenn R. Carroll
ERS-2000-29-ORG

Reasoning with partial Knowledge
László Pólos & Michael T. Hannan
ERS-2000-30-ORG

Applying an Integrated Approach to Vehicle and Crew Scheduling in Practice
Richard Freling, Dennis Huisman & Albert P.M. Wagelmans
ERS-2000-31-LIS

Informants in Organizational Marketing Research: How Many, Who, and How to Aggregate Response?
Gerrit H. van Bruggen, Gary L. Lilien & Manish Kacker
ERS-2000-32-MKT

The Powerful Triangle of Marketing Data, Managerial Judgment, and Marketing Management Support Systems
Gerrit H. van Bruggen, Ale Smidts & Berend Wierenga
ERS-2000-33-MKT

The Strawberry Growth Underneath the Nettle: The Emergence of Entrepreneurs in China
Barbara Krug & Lászlo Pólós
ERS-2000-34-ORG

Consumer Perception and Evaluation of Waiting Time: A Field Experiment
Gerrit Antonides, Peter C. Verhoef & Marcel van Aalst
ERS-2000-35-MKT

Trading Virtual Legacies
Slawomir Magala
ERS-2000-36-ORG

Broker Positions in Task-Specific Knowledge Networks: Effects on Perceived Performance and Role Stressors in an Account Management System
David Dekker, Frans Stokman & Philip Hans Franses
ERS-2000-37-MKT

An NPV and AC analysis of a stochastic inventory system with joint manufacturing and remanufacturing
Erwin van der Laan
ERS-2000-38-LIS

Generalizing Refinement Operators to Learn Prenex Conjunctive Normal Forms
Shan-Hwei Nienhuys-Cheng, Wim Van Laer, Jan N Ramon & Luc De Raedt
ERS-2000-39-LIS

Classification and Target Group Selection bases upon Frequent Patterns
Wim Pijls & Rob Potharst
ERS-2000-40-LIS

New Entrants versus Incumbents in the Emerging On-Line Financial Services Complex
Manuel Hensmans, Frans A.J. van den Bosch & Henk W. Volberda
ERS-2000-41-STR

Modeling Unobserved Consideration Sets for Household Panel Data
Erjen van Nierop, Richard Paap, Bart Bronnenberg, Philip Hans Franses & Michel Wedel
ERS-2000-42-MKT

The Interdependence between Political and Economic Entrepeneurship
ERS-2000-43-ORG
Barbara Krug

Ties that bind: The Emergence of Entrepreneurs in China
Barbara Krug
ERS-2000-44-ORG

What's New about the New Economy? Sources of Growth in the Managed and Entrepreneurial Economies
David B. Audretsch and A. Roy Thurik
ERS-2000-45-STR

Human Resource Management and Performance: Lessons from the Netherlands
Paul Boselie, Jaap Paauwe & Paul Jansen
ERS-2000-46-ORG

Average Costs versus Net Present Value: a Comparison for Multi-Source Inventory Models
Erwin van der Laan & Ruud Teunter
ERS-2000-47-LIS

A Managerial Perspective on the Logic of Increasing Returns
Erik den Hartigh, Fred Langerak & Harry Commandeur
ERS-2000-48-MKT