Scientific Research

Beyond Customer Churn: Generating Personalized Actions to Retain Customers in a Retail Bank by a Recommender System Approach

Michele Gorgoglione, Umberto Panniello

Department of Mechanical and Management Engineering, Polytechnic of Bari, Bari, Italy.
Email: m.gorgoglione@poliba.it, u.panniello@poliba.it

ABSTRACT

Customer churn may be a critical issue for banks. The extant literature on statistical and machine learning for customer churn focuses on the problem of correctly predicting that a customer is about to switch bank, while very rarely considers the problem of generating personalized actions to improve the customer retention rate. However, these decisions are at least as critical as the correct identification of customers at risk. The decision of what actions to deliver to what customers is normally left to managers who can only rely upon their knowledge. By looking at the scientific literature on CRM and personalization, this research proposes a number of models which can be used to generate marketing actions, and shows how to integrate them into a model embracing both the analytical prediction of customer churn and the generation of retention actions. The benefits and risks associated with each approach are discussed. The paper also describes a case of application of a predictive model of customer churn in a retail bank where the analysts have also generated a set of personalized actions to retain customers by using one of the approaches presented in the paper, namely by adapting a recommender system approach to the retention problem.

Keywords: *Customer Churn, Customer Retention, Personalization, Predictive Models, Recommender Systems*

1. Introduction

Retail banks often deal with customer churn. Among the several issues addressed by Customer Relationship Management (CRM), identifying the customers who are about to quit the relationship with a company is one of the most important in the financial services industry. When competition becomes tougher, when laws decrease either the barriers to entry or the customer's switching costs, or when a company aims at strengthening its position in a new market, the issue of retaining customers and avoiding customer churn becomes even more crucial.

In the last decade several computer science scholars have tackled the problem of building accurate models to identify customers at risk in a bank by using statistical, machine learning and data mining approaches [1]. However, much literature only focuses on the problem of correctly predicting that a customer is about to switch. Very rarely the problem of generating personalized actions to improve customer retention rates is considered. The decision of what actions to deliver to what customers is normally left to managers who can only rely upon their

knowledge. However, these decisions are at least as critical as the correct identification of customers at risk. A customer churn prediction model should be integrated with a model to decide what personalized marketing action to deliver to customers; otherwise the benefits of using accurate predictive models would be lost.

The issue addressed by this research is to identify a set of approaches to generating actions to retain customers, once customers at risk are identified and profiled. We believe that this issue is important for both businesses launching customer retention campaigns and researchers aiming at building models of customer churn. From a business perspective, the problem is to support managers in making the right decision on how to define personalized retention actions. In fact, this decision has to consider both organizational aspects, such as control and efficiency, and technological aspects, such as the accuracy of the algorithms. The challenge for managers is to integrate the results of a predictive model of customer churn with the decision of delivering certain marketing actions. The problem is how to associate the right action

Beyond Customer Churn: Generating Personalized Actions to Retain Customers in a
Retail Bank by a Recommender System Approach

91

with the right customer. From a scientific perspective, a comprehensive model of how to generate actions might be helpful to develop more effective data mining and statistical models to support CRM processes.

By looking at both the scientific literature on CRM and personalization, this paper proposes a number of approaches which can be used to generate personalized marketing actions, and discusses benefits and risks associated with each model. Finally, a case of application is described. A retail bank developed a predictive model of customer churn. The model was integrated to a method to generate personalized actions aimed at customer retention. The method is based to the approach followed by Recommender Systems. The results of this model are compared to those obtained by the same company in a formerly launched retention campaign.

2. Prior Research

The problem of customer churn has been analyzed by scholars in the fields of marketing and analytical CRM, by analyzing several aspects. For instance, [2] analyzes the possible payoffs of retention and acquisition strategies depending on the market structure. An analysis of the switching costs is provided by [3]. Customer churn has also been studied by analyzing CRM-related problems such as the effectiveness of loyalty programs [4] and customer satisfaction [5].

Analytical models to predict customer churn have been developed in several areas and industries as well. Marketing models have been deployed to predict customers at risk [6,7] as well as data mining models [8], especially in telecommunication-related industries [9-11]. Specific prediction models have developed also in banking and finance [12,13].

This body of literature have mainly focused on the problem of correctly predicting customer churn. Very rarely the problem of how to generate personalized actions, once customer churn has been modeled, is considered. For instance, [14] states that a good customer churn model should also provide managers with a way to generate personalized actions, although the research focuses only on predicting customer churn. The need to integrate the analysis of customer behavior with the generation of actions in a comprehensive model is maintained by [15], while [16] suggests to include customers found at risk of leaving the company in the organization of events in order to strengthen their relationships with the company.

As the literature on customer churn does not directly face the problem of how to generate actions, and in order to achieve a more general understanding of the problem, we have broadened the focus of the literature analysis to the research on personalization-related areas. The goal is to review the approaches to deliver actions to individual customers. A review of studies on personalization can be found in [17]. The authors consider three main areas, namely computer science and information systems (CS/IS), marketing, management science and economics. It is possible to extract from that review the works directly related to the problem of generating personalized actions.

In the CS/IS area the main streams dealing with actions are Recommender Systems and Web Contents Personalization. In both cases the problem of generating personalized actions is solved by processing individual customer profiles which describe personal preferences. In a recommender system the profile describes a customer's preference by including either a set of ratings to certain products or information on customers' previous purchases [18]. The action is the recommendation of a product and it is generated by analyzing the similarities between the profiles of customers who bought or rated a product and those of customers who did not. These systems are the most common approaches to personalization in real industrial settings ever since Amazon adopted a Recommender System [19]. In the case of Web Content Personalization the user's profile includes statistics on the usage of the Web and the action is generated by associating a content with similarities with the user's profile [20].

Scholars in marketing have dealt with the problem of generating personalized actions by studying competitive personalized promotions [21], customization [22], the personalization of services [23], and recommender systems. In all cases, except that of recommender systems which was reviewed above, the problem is solved by using optimization models that allow the analyst to find the best price and/or product based on individual customers response models. The response models are built by collecting data on customers' previous behavior.

The research in management science [24] and economics has mainly faced the problem of generating personalized actions by studying personalized pricing and mass customization [25]. The first stream falls in the marketing approaches reviewed above. The second stream mainly analyzes the benefits and costs of producing as many products as the number of different customers' individual preferences a company can identify in the market [26]. In these models the action is offering of a vast variety of products. Actions are generated by processing customers' preferences in an aggregated way, not individually, while the choice of the specific product is left to customers.

A few research streams can be added to the review taken from [17] which do not directly deal with personalization, namely Direct Marketing and Knowledge Dis-

covery in Databases (KDD). An extensive literature exists in the area of direct marketing dealing with targeting problems. The largest body of knowledge is related to the process of segmentation, targeting and positioning. It includes several methods to define and select a target market and position an offer, *i.e.* define a specific marketing mix (or any other marketing action) to that specific market. Several scholars in marketing have warned about the fact that a segmentation should be done only aiming at a clear marketing goal [27], *i.e.* segments should be built in order to deliver targeted actions. The research on direct marketing has produced models to define and deliver actions to more granular targets, even single individuals. For instance, "RFM models" represent customers by their *Recency* (*i.e.*, the time since they have made the last purchase), *Frequency* (*i.e.*, how often they purchase), and *Monetary* value (*i.e.*, the average money they have spent) in order to compute the probability to which a customer will make the next purchase by a certain period of time. RFM models have quickly become common in direct marketing and have been extended to give rise to more sophisticated models such as Automatic Interaction Detection Models and Regression Scoring models [28]. In the industrial applications of these models, managers often define a very small set of actions, even one action (e.g. proposing a discounted subscription to a new service), and run the model in order to select a subset of individual customers who represent the optimal target.

Knowledge Discovery in Databases (KDD) is a process to find valid, new, understandable and useful patterns in data [29]. This process can be seen as a possible way to target actions to customers. If the rules derivable from KDD are "actionable", *i.e.* present information on a customer useful to decide "what to do" in order to modify that customer's behavior, then a manager can define marketing actions tailored to that customer [30,31]. The main limitation of these approaches is that either they are not efficient, due to the big amount of discovered rules which need human supervision, or actions are not personal, because when rules are aggregated to improve efficiency the target becomes aggregated as well. Starting from KDD, some attempts were done to define a general method to mining actionable patterns in databases and thus generating personalized actions based on data describing customers' reactions to the company's previous offers [32].

A last remark is useful to highlight that many companies, especially in banking and finance, make use of front office personnel to generate actions tailored to the personal characteristics of customers. In this approaches, managers have to talk to customers, understand their preferences and propose them specific offers. Actions can be

decided a-priori or a-posteriori. In the first case the front office managers select the right customers to whom a certain action can be offered, while in the second actions are decided by the front office managers, who have to be given the necessary degree of autonomy, based on their understanding of customers' needs.

3. Approaches to the Definition of Personalized Actions

By reviewing the literature, five main categories can be identified to classify the approaches to the generation of personalized actions. Each category represents a set of homogeneous approaches which can be used to decide what action should be delivered to what customers in a CRM program. The approaches are reported in **Table 1** and discussed below.

Computational approaches include all those approaches that build a complete model of customers' behavior, actions, and customers' reactions based on information stored in a data set. These approaches can use both data mining [32] and optimization models [21-23]. An example of applications in banking and finance is a retail bank which stores the data related to promotion of stocks, and the reactions of customers who might have purchased those stocks or not. The fundamental condition that enables the adoption of these approaches is the completeness of data. If the data sets do not include, for instance, the reaction of each customer to a certain marketing campaign, then neither a mining algorithm nor an optimization model can be run to generate personalized actions. Computational approaches allow a company to fully automate the generation of actions based on customers' profiles, as no human decision is needed. A limitation is that only marketing actions already launched before can be considered in such approach. The full coverage of customers may be another problem because some customers' reactions may remain unknown (for instance, when a customer does not respond to a survey).

Similarity-based approaches are those used by Recommender Systems [19] and Web content personalization methods [20]. This kind of approach assumes that actions are related to customer preferences, preferences may be inferred by customer profiles, and that either similar customers behave similarly or similar actions cause similar reactions. A "similarity-based" approach does not require to store as much information as a computational approach. Recording customers' preferences is enough, because it is assumed that the unknown preferences of a customer can be derived by identifying the similarity with other customers. However, the twofold condition of applicability of such approaches is that customers' profiles have to represent preferences, and only

actions associated with those preferences can be generated. For instance, a customer who owns multiple credit cards can be classified as a customer who "prefers" using credit cards, whereas the fact that a customer has a mortgage does not necessarily represent a "preference". A "similarity-based" approach is useful to automate the personalization process. An example of this approach in banking is given in Section 5.

Bottom-up approaches include the knowledge discovery methods [31,33] and the use of front office personnel. These approaches consist of two separate steps: 1) profiling customers, 2) deciding proper actions, where the first step has to precede the second step (*i.e.*, actions depend on profiles). They cannot be fully made automatic because only the first step is performed by an algorithm. For this reason these approaches are typically not very efficient. The condition of applicability is that the decision-making effort has not to exceed the company's resources: either the number of customers is low or the number of decision-makers is high. The advantage is that targeting can be very effective because each profile is thoroughly analyzed before generating a proper action. An example of bottom-up approach in banking is the work of a financial advisor who manages a portfolio of customers. The advisor has periodic conversations with a customer, analyzes her needs and proposes tailored financial solutions.

Top-down approaches include the direct marketing approaches [28]. They consist of the same two separate steps typical in bottom-up approaches. However, in this case, the decision of what actions to deliver is made before the definition of customers' profiles and, hence, profiles depend on actions. For instance, a retail bank managers may first decide to offer customers a discount on bank transfers, and then select the target customers by building appropriate profiles. For this reason top-down approaches do not cover the whole customer base: only customers falling in the target profiles are delivered the actions. In this approach, the control on the decision-making process is very high, as the definition of actions is made centrally (it is not delegated to front office personnel as in a bottom-up approach). Profiles can be determined by either running algorithms or asking the front office personnel to identify suitable target customers. The condition of applicability is that actions can be defined before customers' profiles. In some CRM problems, such as reducing customer churn, this condition can be questionable.

Customization approaches are those which offer customers many different options and let the customers choose the suitable one. This approach is typically adopted in mass customization [25]. In this approach actions do not derive from processing customers' profiles but rather from customers' choices. This makes customization different from other personalization approaches [26]. The offer has to be granular enough in order to adopt this approach. For instance, a retail bank can propose customers to choose one among many discounted options (e.g., a discount on bank transfers, credit cards, cash cards, e-transactions, etc.). The association between actions and profiles is performed by the customers instead of the company. For this reason, the control over the process is low and the resulting actions can turn out to be quite expensive. However, the targeting is expected to be quite effective.

4. Which Approach is Best? Discussion of Benefits and Risks

The approaches identified above have characteristics that make them beneficial in certain settings while risky in others. The suitability of each approach should be thoroughly discussed by managers based on the observation of their benefits and risks compared to the company goals and business conditions. Typically, managers have to deal with trade-offs. The main characteristics which can differentiate the five approaches are the following:

- control on the decision-making process;
- automation of the personalization process;
- effectiveness of targeting;
- process efficiency;
- scope of actions;
- coverage of customers.

4.1. Control on the Decision-Making Process

Companies normally want to keep as much control as possible on the definitions of the actions to deliver. Possible reasons are the need of controlling the costs related to marketing actions, and avoiding to leave too much power to front office personnel. In this case, the decision of what offering to customers should be centralized. As an example, a retail bank may want to avoid that the front office personnel makes the decision of what to offer to the customers in their portfolio, because both costs may increase and the bank could be exposed to opportunistic behavior. In this case, the suitable approaches to be used are "computational", "similarity-based" and "top-down". In all these approaches, the decision of what action to deliver is made centrally, by the support of either an algorithm or the strategic orientation. The "bottom-up" and "customization" approaches may be risky because the first may involve the use of front office personnel, the second requires the intervention of customers.

4.2. Automation of the Personalization Process

The need of automating the process of profiling, defining

personalized offers and delivering actions can come from several reasons, such as the need of keeping the cost low and the promptness of CRM processes high. For instance, companies in the financial services industry may want to be prompt and fast in starting a retention program when the churn rate suddenly increases. In this case, automating the process is critical. "Computational" and "similarity-based" approaches are the most suitable when the need of automation is high. In fact, they can be implemented into an algorithm which can be integrated in the company information system. The remaining approaches ("bottom-up", "top-down" and "customization") can only be partially automated, as it was observed in the previous section.

4.3. Effectiveness of Targeting

Companies involved in CRM programs are often concerned in maximizing targeting, *i.e.* to maximize the probability that a customer receives an offer tailored to her own needs. For instance, a retail bank may want to make sure that each customer is offered exactly the product she needs. In these case, the best approaches to define personalized actions are the "computational", "bottom-up" and "customization" ones. In the first approach actions are derived by knowing the customers' response. In the second, actions are defined by analyzing each profile. In the third, customers themselves decide what action they prefer. Targeting is less effective in the "similarity-based" and "top-down" approaches because the customers' need are inferred by similarity or simply not considered when the decision of actions is made. The effectiveness of targeting can be evaluated by the "redemption rate" which measures how many customers react the expected way to a marketing campaign.

4.4. Process Efficiency

Whatever the goal of a CRM program, companies have to keep the costs of CRM as low as possible. Companies may choose the best approach to define personalized actions by considering the level of resources they can allocate to that process. Some approaches requires less resources than others. Particularly, "similarity-based" and "top-down" approaches look suitable to this aim. The algorithms required by the first approach are relatively simple, as it is witnessed by their quick diffusion in industrial settings. The second approach makes the integration between the decision of actions and definition of profiles simpler than other. On the contrary, the algorithms required by a "computational" approach are more complex. A "bottom-up" approach requires either a huge effort in terms of supervision of the knowledge discovery process or a high level of resources in terms of people

involved in the process. The efficiency of a "customization" approach seems to depend on the business settings, as it can be easy to implement but the results can turn out to be expensive to the company. Efficiency can be measured by the number of people involved and/or by time spent on the process.

4.5. Scope of Actions

Companies may want to be creative in the definition of the actions to deliver to customers. The proposed approaches have different impacts on the capability of the company to define a wide spectrum of actions. The only approaches that look suitable to this aim are those in the "bottom-up" category. In fact, it always requires the human supervision, to either infer the needs from the profiles and, in turn, the actions, or talk to each customer, dig into his/her preferences and needs, and finally find out the best offer to deliver. The scope of actions in "computational" and "similarity-based" approaches are limited by the nature of the data. In the former, the only actions which can be offered to customers are those recorded in the databases, for which customers' reactions are already recorded and codified. In the latter, actions have to be related to the variables used to profile customers. In the "top-down" approach, the scope of the actions is usually small because companies define a small set of offers. In the "customization" approach the scope of actions depends on the spectrum of different choices customers are provided with. A potential measure for this variable can be the number of different actions.

4.6. Coverage of Customers

Not all approaches allow a company to fully cover the customer base with personalized actions. The "similarity-based" and "bottom-up" approaches guarantee a full coverage. In the first, actions can be defined by similarity. Even customers for whom the amount of data is very little (e.g., new customers), actions can be derived by the preferences of similar customers. In the second, human decision makers can rely on all their knowledge to define actions starting by profiles. On the contrary, in a "computational" approach the coverage is limited to customers for whom relevant information on actions and reactions have been recorded. In a "top-down" approach, actions are defined independently of profiles, with the consequence that not all profiles can be associated with actions. In a "customization" approach only customers who make an explicit choice are then offered the actions, those who do not want to express any preference remain outside the coverage of the CRM program. The coverage of customers can be measured by the percentage of customers for

whom actions can be derived from profiles.

Table 1 represents each approach in terms of its characteristic, conditions that must be true to adopt that approach, the benefits and risk that should be discussed in order to select the most appropriate approach.

5. A Case of Application

Born as a small local bank, the company went through a significant M&A process over the last few years. It acquired many new customers as a result of this process, currently having about 300 000 customers in total and being a medium-sized Italian retail bank operating at a national level. In 2008 the bank managers had observed a growing churn rate. Although the churn rate was much smaller than the acquisition rate, the cost for acquiring a new customer was much higher than the cost of retaining a customer. Moreover, the cost for the re-acquisition of a lost customer was even higher. These observations made the need of reducing customer churn one of the bank's strategic priorities.

At the end of 2008 a team of managers was involved in a project aimed at both identifying customers at risk of leaving the bank and launching a marketing campaign to retain those customers. By talking to financial advisors and senior managers, the team identified a number of events which normally signal that a customer is about to leave the company. The most important signals were a relevant decrease in the assets, the sale of more than one financial products formerly owned, the cancellation of all the automatic outgoing payments and incoming credits. By analyzing the customer data set, the managers identified the customers at risk by selecting those customers who presented one or more signals of imminent attrition.

They were ordered by the number of signals of attrition in a ranking from high to medium risk.

Once the customers were identified, the team had to carefully face the problem of what actions should be part of a marketing campaign. A fairly straightforward choice would be leaving the decision to the personal financial advisors, who are sales reps with a portfolio of about 150 (on average) customers to manage. Since they know each customer personally, they could decided the best offer to retain each customer. However, the bank top management decided that the actions to deliver to customer had to be identified centrally. The reason was threefold. Firstly, not all customers had a direct and strong relationship with their personal advisors. Some preferred to interact with the bank via the home banking system or other virtual channels. Secondly, the high number of personal advisors, about 2000, would make the process quite expensive. The auditing process would be slow and complex. Thirdly, the top managers wanted to keep a strong control over the whole customer relationship management process. The problem was too crucial to be left to the initiative of front office personnel.

Therefore, in order to generate appropriate retention actions, the managers had several meetings with senior managers and sales/marketing managers. A list of few marketing actions was generated, such as offering the customer a credit card for free or a discount on the annual fee. The list was communicated to all the personal financial advisors. They were told to pick one or more actions among those in the list for each customer identified at risk of leaving the bank, depending on the customer's profile. The actions had to be suitable to each specific customer and picked based on the advisors'

Table 1. Approaches to the definition of personalized retention actions.

Approach	Characteristics	Conditions	Benefits	Risks
Computational	Complete model of preferences, actions and reactions	The data set includes information on prior marketing actions and customers' responses	• Control • Automation • Targeting	• Limited scope of actions • Customer coverage
Similarity-based	Comparison between customer profiles	Actions are related to customer preferences, and preferences may be inferred by customer profiles	• Control • Automation • Efficiency • Coverage	• Limited scope of actions
Bottom-up	The definition of customers' profiles precedes the definition of actions	The supervision effort has to be affordable (in terms of number of profiles and resources)	• Targeting • Scope of actions • Coverage	• Low efficiency • Lack of control
Top-down	The definition of customers' profiles follows the definition of actions	Actions can be defined independently of customers' profiles	• Control • Efficiency	• Lack of targeting • Limited coverage of customers
Customization	Customers are free to choose the appropriate action	The offer has to be granular enough	• Targeting	• Low efficiency • Lack of control

JILSA

knowledge of the customer's behavior.

As a result, a redemption rate of 78.5% was observed, meaning that 1062 customers out of 1353 contacted customers responded positively to the marketing action proposed and did not leave the bank, thus reducing the overall churn rate.

Although the results were not impressive, they were fairly aligned with the managers' expectations. In fact, the manager themselves were aware of some flaws in the first project. First of all, actions were independent of customers' profiles. The decision of what to offer to customers at risk was made upstream, based on previous experiences rather than a deep knowledge of each customer. Therefore, a very small number of customers would receive an offer tailored to their actual needs. Another flaw of the method was the very limited possibility to automate the process. While the identification of customers at risk could be implemented on an information system, the decision-making process through which to identify retention actions could not.

Encouraged by the results and the commitment of the bank executives, the team decided to launch a new, improved project of customer retention. The main goal would be improving targeting and the possibility of automating the process.

5.1. Predictive Model of Customer Churn

The managers decided to collaborate with a group of analysts, including this paper's authors. The project followed the Cross Industry Standard Process for Data Mining (CrispDM) methodology [34]. The first step was aimed to codify the managerial knowledge and beliefs, to define organizational and business constraints and to identify the sources of useful information. This was helpful to guide the analysts in the definition of the variables that would be included in the model. For instance, in the managers' experience the more the products owned by a customer the higher the probability that she would be loyal, while customers owning only one product had the highest probability of leaving the company. The products which made customers loyal were typically financial products such as loans, stocks and bonds.

In the next step, the analysts discussed with the managers the nature of the data sets to use based on the first step output. The data warehouse service was outsourced by the bank. Therefore, any data set had to be explicitly asked to the outsourcer which managed the data storage. Firstly, the team decided to focus only on customers in the "basic" segment with at least a charged bank account. The "basic" segment was the one with the highest number of customers and the highest churn rate. Customers in higher segments shall not be included in the analysis.

Secondly, the group decided not to include the data related to customers in the branch offices most recently acquired from other banks, in order to keep the data homogeneous. Thirdly, the data would be referred to the last two years. The reason was that the bank had older data only for some customers. The data available on customers coming from other banks recently acquired by M&A processes was not older than two years. Moreover, a technical constraint in the storage service made the retrieval of data referred to years preceding the last two very slow. Finally, the data would include demographical information on customers, such as age and job, the history of the relationship between customers and bank (e.g., when an account was open or when a financial product was purchased and/or sold), the products owned by customers (e.g., credit cards, cash cards, stocks, insurances, etc.), the operations made on the account. Operations were classified and divided into subcategories: 1) payments by card, bank transfers and cheques (incoming and outgoing), 2) payments to utility services, 3) salaries and similar incomes credited on the account, 4) other operations. Other information for each customer were also included in the data set, such as the name of the customer's personal financial advisor, the branch office the customer belongs.

Once the group was provided with a data set, a preliminary analysis was run in order to check the correctness of the data. The acquisition and analysis proceeded in an iterative way, until a complete and correct set of data was provided. However, the data on operations on customers' accounts was available only for the last year. For this reason that specific data was not used in the subsequent modeling phase. The initial dataset included around 200 000 customer, 300 000 records and 100 variables.

The next step was aimed to clean and modify the data according to the analysis' goals and method. Some records were deleted while others were re-coded, and several variables were transformed starting from the initial data set. The records referred to users older than 70 years were dropped from the database (customers older than 70 could leave the bank for "natural" reasons more than for a real choice). The records referred to customers with more than one account were merged into one overall record, in order to obtain a data set where each customer corresponded to one record. Several variables were transformed. Some were discretized (e.g., length of the relationship between customer and bank). For some variables the analysts computed statistics such as average, maximum and minimum value (e.g., operations made on the account). Mistakes were cleaned (e.g., double codes for the same value of a variable) and some records with

Beyond Customer Churn: Generating Personalized Actions to Retain Customers in a
Retail Bank by a Recommender System Approach

97

too many missing values were deleted.

The final number of customers in the data set was 203 196. More than 200 variables were included in the new data set, although only a small subset would be used in the subsequent modeling phase. The depth of historical data owned by the bank was two years, namely 2007 and 2008. During the analysis which was performed in the first months of 2009, the bank continued to gather similar data in order to apply the model to a new set of data.

The core step of the project was aimed to build a predictive model of customer churn and test it on actual data. After several experiments, a relatively small subset of 27 variables, taken from the initial set of 200, was selected to build the model. The set included the length of the relationship between customer and bank, the customer's place of residence, the type of card owned (e.g., cash, credit or pre-paid card), the type of financial products and services owned (e.g., stocks, loans, insurances) and their quantity and value, the type of account owned (corresponding to a certain set of discounted options on the account), the amount of payments and incomes automatically debited and credited, respectively, on the account (see **Table 2**).

After several trials, the algorithm used to build the predictive model was J4.8. The training and validation sets included all data referred to 2007 and 2008. Each point was represented by the figure at the end of each semester in 2007 and 2008, therefore 4 vectors of 27 variables for each customer were used at most to train the model. A ten-fold cross-validation method was used. This data included information about 11,000 customers who had abandoned the bank between 2007 and 2008 and about 180,000 customers who had remained loyal in the same period. An additional test was performed by applying the model, once trained and validated, to the data related to the first semester of 2009. This data set was gathered while the analysts were building the model and the test consisted of applying the selected model, with the selected settings, to the actual data in order to predict the churner users.

Two predictive models were built by the analysts and provided to the bank managers who would select one, based on some managerial implications. The models predicted whether a customer would leave the company or not, based on the data. In fact, because of the limited temporal depth, it was not possible to train the model to predict customer churn within a specific period of time. The output of the models were two lists of customers with a predicted binary value ("will leave" or "will be loyal") and a corresponding score. The first model selected 11,070 potential churners among the active customers with an accuracy equal to 87.35% and a true posi-

Table 2. Variables used in the modeling phase.

Variable	Type
Length of customer-bank relationship	numerical[1]
Place of residence	nominal
Type of card owned	nominal[2]
Maximum value of a customer's transaction	numerical
Type of transaction corresponding to the maximum value	nominal
Maximum quantity of a product purchased	numerical
Type of product corresponding to the maximum quantity purchased	nominal
Type of account owned by the customer	nominal[3]
People sharing the same account with the customer	binary
Products owned by the customer	nominal[4]
Automatic payment of utility services	binary
Automatic credit of incomes on the account	binary

[1]All the numerical variables were discretized; [2]split in 4 binary variables; [3]split in 3 binary variables; [4]split in 11 binary variables.

tive rate equal to 0.916. The second model selected 32 938 potential churners among the active customers with an accuracy equal to 67.34% and a true positive rate equal to 0.994. The potential "churners" identified by the first model were a subset of those selected by the second model, except 282 customers (0.14% of the whole customer base).

The bank's managers decided to select the second model. The decision was based on the fact that the cost of a false negative was considered substantially higher than the cost of a false positive. The managers preferred to contact more customers, even loyal customers misclassified as potential "churners", than risking not to identify actual customers at risk. Moreover, the second model performed significantly better than the first, once tested on the first semester of 2009, as the second identified 96.89% of customer who actually left the bank, whereas the percentage of re-identification in the first model was 79.33%.

5.2. Selection of a Suitable Approach to the Generation of Personalized Actions

The next problem was to integrate churn prediction with the decision of which marketing actions should be delivered to which customer. The analysts proposed all the five approaches presented in the section above to generate personalized actions for those customers predicted as potential "churners". Firstly, the team discussed the con-

ditions of applicability of each approach. Secondly, the main managerial issues were ranked by priority in order to discuss benefits and risks.

The only condition to discard, among those listed in **Table 1**, was the one related to the "computational" approach. Prior actions and customers' response data had never been stored in the dataset by the bank. Given the absence of this kind of data, a "computational" approach could not be adopted. All the other approaches could be potentially adopted. Actions could be inferred by customer profiles by a "similarity-based" approach, at least partially, because they included some behavioral data (products owned). The bank could afford the supervision of effort in a "bottom-up" approach if the process was delegated to front office personnel. Actions could be defined independently of customers' profiles, in a "top-down" approach, because managers had some experience on prior campaigns. Finally, the offer was granular enough to let the customers choose in a relatively wide set of alternatives, thus adopting a "customization" approach.

The discussion of benefits and risks was crucial to select the most appropriate approach. The two main issues raised by the managers were ranked as (1) keeping high control on the marketing process, (2) achieving better targeting with respect to prior marketing campaigns. According to the first issue, the team excluded the adoption of both a "bottom-up" and a "customization" approaches. In fact, given the high number of customers predicted at risk, a "bottom-up" approach would be feasible only by delegating the task of revising each profile and deciding the corresponding actions to the front office personnel. This would entail a substantial loss in control. The same would happen by letting customers choose among different optional actions in a "customization" approach. The choice between a "similarity-based" and a "top-down" approach was eventually made based on the second issue. In fact, the bank's prior marketing cam-

paigns had been carried out by defining non-personalized actions, and the goal was just to improve targeting. A "top-down" approach would represent too small a step towards personalization, while a "similarity-based" approach would allow the bank to achieve a better targeting. **Figure 1** reports a graphical representation of the process followed to select the most suitable approach to the definition of personalized actions.

5.3. Generating Retention Actions by a Recommender System

The most common practical applications of the "similarity-based" approach are Recommender Systems (RS). Given a set of customers and a set of products or items, a RS predicts the unknown utility of an item for a customer. If item j is predicted to be of a high utility for customer i, then the system delivers the marketing action of recommending that customer to purchase that item. The action is personalized because the set of items recommended to customers is different for each customer.

More formally, a RS deals with two types of entities: users (e.g., customers) and items (e.g., products). In the transaction-based RSs, the utility of an item for a customer is measured by a Boolean variable indicating if the user owns a particular item or not, or with the purchasing frequency of an item, or with the usage frequency of that item [35,36]. Based on the known values of utility, a RS tries to estimate the utility of the yet unseen items for each customer. In other words, a RS can be viewed as the rating function R that maps each user/item pair to a particular utility value [18]:

$$R: Users \times Items \rightarrow Utilities \qquad (1)$$

One of the main tasks of a RS is to make this function total by estimating the unknown utilities. The estimation is done by using a similarity function, which can be designed by following one of the following three approaches [18]:
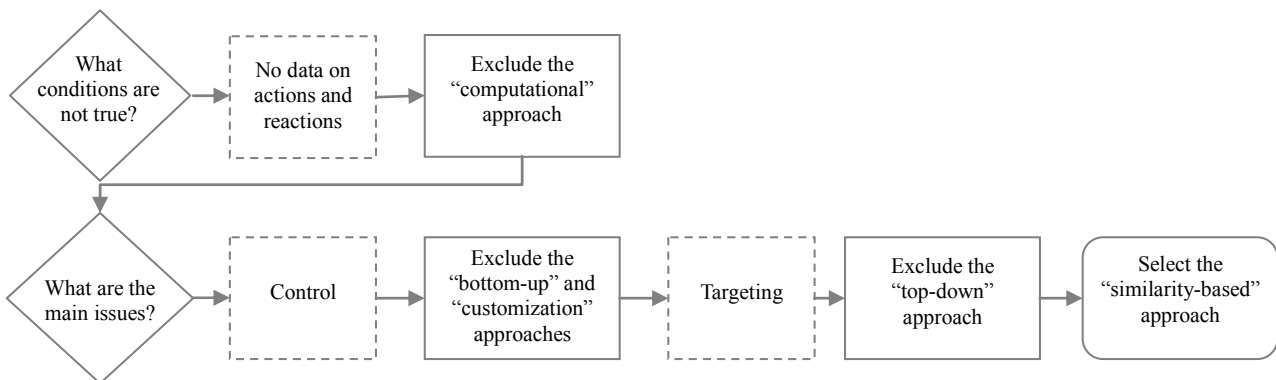
Figure 1. Representation of the decision-making process.

- Content-based recommendations: the user is recommended items similar to the ones the user preferred in the past;
- Collaborative-filtering recommendations: the user is recommended items that people with similar tastes and preferences liked in the past;
- Hybrid approaches: they combine collaborative-filtering and content-based methods.

This general approach had to be adapted to the problem of generating retention actions to customers at risk of churning. Three hypotheses were identified in order to consider the approach feasible: 1) recommending certain products to customers can improve retention; 2) the products able to improve customer retention are not the same for all customers and can be identified by looking at the behavior of loyal customers; 3) similar customers behave similarly and have similar preferences.

While the third hypothesis has much support in the marketing literature, the first two were significantly supported by the management expertise and partially by empirical evidences. In fact, the bank's managers were confident that certain products, such as bank's stocks, insurance products and loans, directly contribute to improve customers' loyalty. These beliefs had emerged in the first project step in the statements of several managers interviewed by the analysts. The managers were also persuaded that the products which can improve retention are not the same for all customers, as each customer may have a different sensitivity to different products. These differences come from several causes, such has different personality traits and different life cycle stages. Because of this variety of behavior, the statistical analyses only partially validated these beliefs. Customers had been split into loyal and those who had abandoned the bank, although some differences could be observed between the two groups in the distribution of the products owned by customers, they were statistically significant only at a low probability.

As a consequence of these observations, the team of analysts and managers decided to adopt a collaborative filtering approach to recommendations, where each customer at risk would be recommended products that loyal customers with similar tastes and preferences liked in the past.

Each user was defined as a vector containing both demographic and transactional data. That kind of information was used because two users can be defined as similar only if they are similar in terms of preferences (e.g., products and services owned by customer, quantity and amount of transactions) and other behavioral and demographical variables (e.g., length of relationship with the

bank, place of residence).

According to the standard collaborative filtering approach, the "neighborhood", *i.e.* the set of customers z with tastes and preferences similar to customer i, was formed by using the cosine similarity, which is given by the following formula:

$$sim(i,z) = \cos(i,z) = \frac{i \cdot z}{\|i\|_2 \times \|z\|_2} = \frac{\sum_{s \in S_{iz}} r_{i,s} r_{z,s}}{\sqrt{\sum_{s \in S_{iz}} r_{i,s}^2} \sqrt{\sum_{s \in S_{iz}} r_{z,s}^2}} (2)$$

where $r_{i,s}$ and $r_{z,s}$ are the ratings of item s assigned by user i and user z respectively. $S_{iz}=\{s \in Items/r_{i,s} \neq \varnothing \wedge r_{z,s} \neq \varnothing\}$ is the set of all items co-rated by both user i and user z, and $i \cdot z$ denotes the dot-product between the vectors i and z. In this specific case, the customer i belongs to the set of customers predicted as churner, whereas the customer z belongs to the set of loyal customers Z. Only the customers predicted as loyal with a probability higher than 90% were considered in Z. The strong hypothesis of this approach is that similar customers behave similarly, so a customer predicted to leave the company can be retained if she is recommended products characterizing the loyal customers with similar tastes and preferences.

Therefore the main difference between a standard recommender system approach and the system used in this case is the way the neighborhood is formed. While in the former, the neighborhood include all the customers similar to the customer to whom an action has to be delivered, in the latter the customers included in the neighborhood were taken from a subset of all customers, namely the most loyal customers. The final output of the recommendation algorithm was a list of products that should be suggested to each customer at risk.

As a result, a redemption rate of 81.6% was observed, meaning that 1313 customers out of 1609 contacted customers, responded positively to the marketing action proposed and did not leave the bank, thus reducing the overall churn rate.

5.4. Discussion of Results

The method used in the first project can be classified as a hybrid of a "top-down" and a "bottom-up" approaches. In fact, on the one hand the team of managers decided what actions should be delivered to customers independently of the profiles of customers at risk. The decision was made by talking to senior managers and was based on a general knowledge of the way customer reacted to marketing campaigns in the past. On the other hand, the team asked the front office personnel (the personal financial advisors) to select the most appropriate action

JILSA

among those identified by senior managers. This selection would be based on each customer's profile.

The flaws of this method mainly came from those associated with a "top-down" approach. In fact, the front office personnel (the "bottom-up" component of the process) was involved only downstream and had a small role. In fact, the main flaws were the lack of targeting and the impossibility of automating the process. Despite the flaws, the method had some benefits, again those associated with a "top-down" approach. Firstly, the bank had the control over the process, because the decisions were made centrally. Secondly, the process was relatively efficient compared to other methods, because the main resource spent was the time of junior and senior managers. The second project should then improve targeting and automation without decreasing control and efficiency.

The method used in the second project can be classified as a "similarity-based" approach. The possibility of defining actions directly by computing similarity was given by the fact that some of the variables used to define customers' profile were "actionable". Particularly, the ownership of products (cards and financial products, see **Table 2**) could be used to derive actions, *i.e.* the offer of products not owned by the target customer but owned by loyal customers similar to the target customer.

The results reached by the second project were consistent to the expectations, at least partially. Control and efficiency did not decrease. In fact, the decisions were made again centrally by the team. The involvement of the front office personnel was no longer necessary, not even to select the most appropriate action, because actions were already tailored to each customer's profile. The time spent to talk to senior managers was approximately the same. However, a fair comparison of the efficiency is not possible, because in the second project the team used the knowledge gained in the first to select the appropriate variables to build the customers' profiles and draw actions. Therefore, the time spent with senior managers was very small, because they only had to approve the set of actions defined by the team.

Automation remarkably improved. At the end of the project, the variables to build the customers' profiles were defined, the data mining model was set, and the algorithm to compute the actions by similarity was codified. At this stage, the team had only to implement a procedure to get the right information from the databases and run the algorithms periodically.

The level of targeting only slightly increased. Trying to explain this result is helpful to highlight the limits of a "similarity-based" approach. Two main reasons can explain the observation.

First of all, a customer's decision of either leaving a company or remaining loyal is affected by many factors. Some are related to the specific products the company offers to customers. For instance, a customer may find more convenient the offer of a competitor and decides to switch. Some are related to the relationship between customer and company. For instance, a customer may feel like she is not treated the right way or the relationship is not as intense as it should be in order to curb the decision of leaving. Other factors pertain to customers' personal conditions and situations, such as when a customer decides to quit because of a weak financial situation suggesting to save money. Given this behavioral complexity, a "similarity-based" approach makes the very simple hypothesis that the decision of remain loyal only depends on improving the company offerings.

The second reason is that a "similarity-based" approach is intrinsically characterized by a low "scope of actions", as it was defined in Section 4, and this ends up to affect the level of targeting. The number of different actions that such approach can identify is limited by the nature of the data. In the case described above, the actions were associated with the variables describing whether the customer owns a product (cards or financial products). Therefore, the company can only offer a customer one or more products she does not own yet, chosen among those owned by similar customers predicted as loyal. This generated a very limited set of feasible actions, far away below the number of possible combinations of the set of products. Because of this intrinsic limitation, the number of different actions the company could offer was so lower than the number of customers, that many customers received the same offer given to other customers. Actually, the set of actions coming from the application of the "similarity-based" method was very similar to the actions the managers had proposed in the first marketing campaign. The consequence of this limited targeting can explain the fact that several customers did not react positively to the marketing action, so reducing the overall redemption rate. The slight increase in the redemption rate was probably due to a better predictive model of customer churn rather than to a better method used to associate the actions with the profiles. In fact, while the first method was based on a set of "alert signals", the second had learnt a customer behavior model. However, the results obtained by the "similarity-based method" still remain interesting, because they demonstrate the applicability of the method in a real industrial setting.

6. Conclusions

The issue tackled by this paper is how to integrate the problem of predicting customer churn through analytical

models with the problem of generating personalized actions to retain customers. The two problems have been treated separately by research so far, as the extant literature witnesses. On the one hand, research has developed accurate statistical and machine learning models to associate each customer's profile with a churn score. On the other, the decision of what personalized action to deliver to each customer has remained a problem for managers.

This paper presented a list of possible approaches to generate personalized actions and discussed the relationship between the generation of actions and the customer profiling. Benefits and risks associated to each approach are also discussed. A case of application was described to support the theoretical discussion. In a retail bank, a team of analysts developed a predictive model of customer churn and integrated it with a method to generate personalized actions to retain customers. Actions were generated by using a "similarity-based" approach. The specific approach was selected by firstly discussing the conditions of applicability of each approach and secondly the main business priorities. This allowed the analysts to discard the unfeasible approaches, and select the most appropriate one. In the business case described, the personalized actions for customer retention were generated by adapting the algorithm of a recommender system. The main difference between a standard recommendation engine and the one used in the application case is the way the users neighborhood was formed, namely by computing the similarity between customers predicted at risk of churn and loyal customers.

The results were aligned with the expectations, at least partially. In fact, the method was expected to improve targeting and automation, without decreasing control and efficiency. Control and efficiency remained the same, while automation remarkably increased. The level of targeting did not increase as much as expected, because of two main reasons. The first is that a complex behavior, the customer's decision of leaving or remaining loyal, is reduced to a much simpler hypothesis, namely the fact that the decision only depends on what products the company offers. The second is that the scope of actions in a "similarity-based" approach is quite limited, and this intrinsically reduces the level of targeting. Further research will be done to reduce these limits and to extend the comparison to other approaches outlined in the paper.

REFERENCES

[1] J. Hadden, A. Tiwari, R. Roy and D. Ruta, "Computer Assisted Customer Churn Management: State-of-the-Art and Future Trends," *Computers and Operations Research*, Vol. 34, No. 10, 2005, pp. 2902-2917. doi:10.1016/j.cor.2005.11.007

[2] N. B. Syam and J. D. Hess, "Acquisition Versus Retention: Competitive Customer Relationship Management," *Working Paper*, University of Houston, Houston, 2006.

[3] P.-Y. Chen and L. M. Hitt, "Measuring Switching Costs and the Determinants of Customer Retention in Internet-Enabled Business: A Study of the Online Brokerage Industry," *Information Systems Research*, Vol. 13, No. 3, 2002, pp. 255-274. doi:10.1287/isre.13.3.255.78

[4] R. N. Bolton, P. K. Kannan and M. D. Bramlett, "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," *Journal of the Academy of Marketing Science*, Vol. 28, No. 1, 2000, pp. 95-108. doi:10.1177/0092070300281009

[5] T. Hennig-Thurau and A. Klee, "The Impact of Customer Satisfaction and Relationship Quality on Customer Retention: A Critical Reassessment and Model Development," *Psychology and Marketing*, Vol. 14, No. 8, 1998, pp. 737-764.

[6] S. A. Neslin, S. Gupta, W. Kamakura, J. Lu and C. H. Mason, "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models," *Journal of Marketing Research*, Vol. 43, No. 2, 2006, pp. 204-211. doi:10.1509/jmkr.43.2.204

[7] Z. Jamal and R. E. Bucklin, "Improving the Diagnosis and Prediction of Customer Churn: A Heterogeneous Modeling Approach," J*ournal of Interactive Marketing*, Vol. 20, No. 3-4, 2006, pp. 16-29. doi:10.1002/dir.20064

[8] J. Hadden, A. Tiwari, R. Roy and D. Ruta, "Churn Prediction: Does Technology Matter?" *International Journal of Intelligent Technology*, Vol. 1, No. 1, 2006, pp. 104-110.

[9] S. V. Nath and S. Behara, "Customer Churn Analysis in the Wireless Industry: A Data Mining Approach," *Proceedings of the 34th Meeting of the Decision Sciences Institute*, Washington, 22-25 November 2003, pp. 505-510.

[10] L. Bin, S. Peiji and L. Juan, "Customer Churn Prediction Based on the Decision Tree in Personal Handyphone System Service," *Proceedings of International Conference on Service Systems and Service Management*, Chengdu, 9-11 June 2007, pp. 1-5.

[11] J. Lu and O. Park, "Modeling Customer Lifetime Value Using Survival Analysis—An Application in the Telecommunications Industry," *Data Mining Techniques*, SAS Users Group International, Vol. 28, 2003, pp. 120-128.

[12] T. Mutanen, "Customer Churn Analysis: A Case Study," *Research Report*, No. VTT-R-01184-06, 2008.

[13] D. V. D. Poel and B. Larivière, "Customer Attrition Analysis for Financial Services Using Proportional Hazard Models," *European journal of Operational Research*, Vol. 157, No. 1, 2004, pp. 196-217. doi:10.1016/S0377-2217(03)00069-9

[14] S. Y. Hung, D. C. Yen and H. Y. Wang, "Applying Data Mining to Telecom Churn Management," *Expert Systems with Applications*, Vol. 31, No. 3, 2006, pp. 515-524.

doi:10.1016/j.eswa.2005.09.080

[15] L. S. Yang and C. Chiu, "Knowledge Discovery on Customer Churn Prediction," *Proceedings of the* 10*th World Scientific and Engineering Academy and Society*: *International Conference on Applied Mathematics Dallas*, Texas, 1-3 November 2006, pp. 523-528.

[16] C. B. Bhattacharya, "When Customers are Members: Customer Retention in Paid Membership Contexts," *Journal of Marketing Science*, Vol. 26, No. 1, 1998, pp. 31-44.

[17] G. Adomavicius, Z. Huang and A. Tuzhilin, "Personalization and Recommender Systems," *Tutorials in Operations Research*, INFORMS, Charlotte, 2008.

[18] A. Tuzhilin and G. Adomavicius, "Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, No. 6, 2005, pp. 734-749.

[19] G. Linden, B. Smith and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing*, Vol. 7, No. 1, 2003, pp. 76-80. doi:10.1109/MIC.2003.1167344

[20] B. Mobasher, R. Cooley and J. Srivastava, "Automatic Personalization Based on Web Usage Mining," *Communications of the Association for Computing Machinery*, Vol. 43, No. 8, 2000, pp. 142-151. doi:10.1145/345124.345169

[21] J. Zhang and M. Wedel, "The Effectiveness of Customized Promotions in Online and Offline Stores," *Journal of Marketing Research*, Vol. 46, No. 2, 2009, pp. 190-206. doi:10.1509/jmkr.46.2.190

[22] A. Ansari and C. F. Mela, "E-Customization," *Journal of marketing research*, Vol. 40, No. 2, 2003, pp. 131-145. doi:10.1509/jmkr.40.2.131.19224

[23] J. Pancras and K. Sudhir, "Optimal Marketing Strategies for a Customer Data Intermediary," *Journal of Marketing Research*, Vol. 44, No, 4, 2007, pp. 560-578. doi:10.1509/jmkr.44.4.560

[24] B. P. S. Murthi and S. Sarkar, "The Role of the Management Sciences in Research on Personalization," *Management Science*, Vol. 49, No. 10, 2003, pp. 1344-1362. doi:10.1287/mnsc.49.10.1344.17313

[25] A. Kumar, "From Mass Customization to Mass Personalization: A Strategic Transformation," *International Journal of Flexible Manufacturing Systems*, Vol. 19, No. 4, 2007, pp. 533-547. doi:10.1007/s10696-008-9048-6

[26] N. Arora, X. Dreze, A. Ghose, J. Hess, R. Iyengar, B. Jing, Y. Joshi, V. Kumar, N. Lurie, S. Neslin, S. Sajeesh,

M. Su, N. Syam, J. Thomas and Z. Zhang, "Putting One-to-One Marketing to Work: Personalization, Customization and Choice," *Marketing Letters*, Vol. 19, No. 3, 2008, pp. 305-321. doi:10.1007/s11002-008-9056-z

[27] J. H. Myers, "Segmentation and Positioning for Strategic Marketing Decisions," South-Western Educational Pub, Cincinnati, 1996.

[28] W. J. Reinartz and R. Venkatesan, "Decision Models for Customer Relationship Management," In: B. Wierenga, Ed., *Handbook of Marketing Decision Models*, Springer Netherlands, Dordrecht, 2009.

[29] U. Fayyad, G. Piatetsky-Shapiro, P. Smith and R. Uthurusamy, "Advances in Knowledge Discovery and Data Mining," The Massachusetts Institute of Technology Press, Massachusetts, 1996.

[30] A. Silberschatz and A. Tuzhilin, "On Subjective Measures of Interestingness in Knowledge Discovery," *Proceedings of the* 1*st International Conference on Knowledge Discovery and Data*, Montreal, 20-21 August 1995, pp. 275-281.

[31] B. Mobasher, B. Berendt and M. Spiliopoulou, "Knowledge Discovery and Data for Personalization," *Tutorial at the* 12*th European Conference on Machine Learning*, Freiburg, 5-7 September 2001.

[32] K. Wang and Y. Jiang, "Mining Actionable Patterns by Role Models," *IEEE International Conference on Data Engineering*, Atlanta, 3-7 April 2006, pp. 16-25.

[33] A. Tuzhilin and G. Adomavicius, "Export-Driven Validation of Rule-Based User Models in Personalization Applications," *Data Mining and Knowledge Discovery*, Vol. 5, 2001, pp. 33-58. doi:10.1023/A:1009839827683

[34] C. Shearer, "The CRISP-DM Model: The New Blueprint for Data Mining," *Journal of Data Warehousing*, Vol. 5, No. 4, 2000, pp. 13-22.

[35] Z. Huang, W. Chung and H. Chen, "A Graph Model for E-Commerce Recommender Systems," *Journal of the American Society for Information Science and Technology* ,Vol. 55, No. 3, 2004, pp. 259-274. doi:10.1002/asi.10372

[36] Z. Huang, X. Li and H. Chen, "Link Prediction Approach to Collaborative Filtering," *Proceedings of the 5th ACM/ IEEE-CS Joint Conference on Digital Libraries*, Denver, 7-11 June 2005, pp. 14-142.