Promoting
Credit Card Usage
by Mining Transaction Data

Peter van der Putten[1]

[1] Leiden Institute of Advanced Computer Science
and Sentient Machine Research, Amsterdam
pvdputten@smr.nl

Abstract. In this article we sketch an application specific datamining scenario for promoting credit card usage, using real world transaction data. We identify two different business goals: finding new credit card customers and upgrading existing cardholders. Analyzing the data, we demonstrate that credit card usage is an important trend for the bank. Both business goals are approached with predictive and descriptive datamining methods. The main datamining result is a model that predicts scores for credit card ownership. We conclude with giving suggestions for exploiting and improving the models.

1. Introduction and Motivation

Today datamining, the continuous analysis of customer data, is widely recognized as an important tool to create learning relationships with customers. However to cross the chasm to widespread, mainstream application (Moore 1998), it will be necessary to develop and promote target group specific application scenarios rather than stressing the benefits of datamining in general.

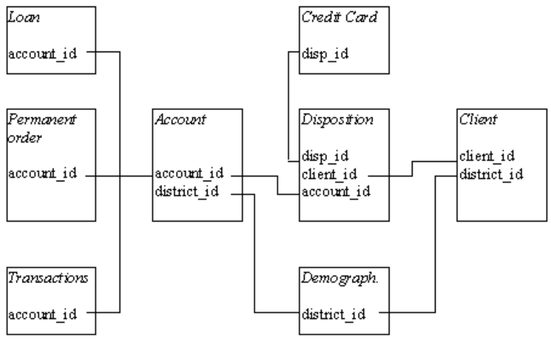
The goal of this article is twofold. First, we strive to demonstrate some of the potential of knowledge discovery methods for detecting consumption patterns in bank transaction data[1]. Second, although this is just a single case, we try to indicate and illustrate some of the universal problems one can encounter in this context, and suggest possible approaches to these problems. Although a rather straightforward business objective was chosen - promote credit card usage - and the datamining problem was constrained by the relative small amount of customer attributes that were available, there are still a great number choices that can be made when detailing the datamining objective, preparing the data, evaluating the discovered knowledge and models and deploying the results.

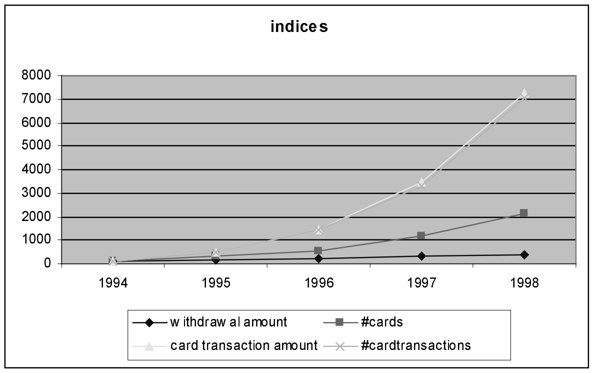
2. The Business Case: Marketing a Credit Card Product

The business case was based on real world bank data, which was made available through the "Discovery Challenge" competition, organized as part of the the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases in Prague (September 1999). The database contains information on 5369 clients, over 1 million transactions, socio-demographic survey data grouped by districts and regions, accounts, permanent orders, granted loans and issued credit cards (see figure 1). Some initial exploratory data analysis revealed a number of interesting business problems which could be approached with datamining, such as loan approval, life time value estimation and retail network planning. In the end we decided to focus on the business goal to promote credit card usage. Credit cards are important for a bank for a number of reasons. First, it is yet another product the bank can cross sell to existing clients and strengthen the relationship. The more bank products a client owns, the less likely it is that the client will switch to another bank. Second, the bank receives a yearly fee and a certain percentage of all purchases.

The particular importance for this bank is not obvious at first sight, but lies more hidden in the data. Out of all withdrawal transactions only 1.2% is a credit card transaction, which corresponds to 0.6% of all withdrawal amounts. In

[1] More datamining cases for direct marketing purposes can be found in van der Putten (1999)





1993, the first year for which there is data, there was only 1 credit card customer! However, at the end of 1998, 18% of all customers own a credit card.
To evaluate if this an important trend, we must compare it to other key indicators; for instance in 1998 there are almost three times more active accounts than in 1994, and the total withdrawal amount was increased by a factor of 3.5. The potential becomes clear if we compute indices for total withdrawal amount, number of credit cards, total credit card withdrawal amount and number of credit card transactions (figure 2). The credit card withdrawal amount increases much faster than the total withdrawal amount. This is due to an increase in number of credit cards and number of transactions per credit card (3.1 to 10.2 transactions per credit card per year); the average transaction amount stays more or less the same. Summarizing: while current use of credit cards is still relatively limited, this situation will probably change rapidly.
So in general it would be advisable to set a business goal to promote credit card use even further. A good strategy to follow would be to 1) promote credit card ownership within the existing customer base (cross selling) and 2) stimulate credit card owners to use their card more often (upgrading).

3 Datamining Efforts

The business goals, cross selling and upgrading, were both approached with predictive and descriptive datamining methods. This section describes the original datamining goals, gives an overview of the data preparation actions that were taken and provides a short description of the mining algorithms and methodology.

3.1 Datamining goals

The goals for descriptive datamining were rather limited given the kind of customer attributes that were available. We planned to build customer profiles for credit card customers (cross selling) and differentiate GoldCard clients from ClassicCard clients (standard card) , and JuniorCard vs Classic Card clients (upgrading). To build the profiles we used straightforward univariate deviation detection techniques which extract and order the most typical differences between two groups.
We planned to construct prediction models for cross-selling and for upgrading. The goal of the upgrading model was to predict how much an existing credit card client would use the card, so that the client potential could be contrasted with the actual consumption behavior. Preliminary modeling showed that the quality of the extracted dataset was too poor to predict credit card value, so we did no further detailed analysis within this context.
However, these models still proved to be useful for cross selling analysis. The goal in this case is to construct a model to predict the potential for a client to own a credit card. The approach was to identify customers who look like credit

card owners, but don't have a credit card. Note there is a subtle difference with identifying clients which are about to get a credit card.

3.2 Data preparation

Before discussing the various preprocessing steps taken to create the target attributes and a client-oriented consumption history, we would like to point out some underlying assumptions and risks. In real world applications these kinds of situations are often encountered.
When using datamining results to identify prospects for credit cards or credit card upgrades within the existing customer base the assumption is made that the data is a representative sample of the entire customerbase. This probably not the case, given the small amount of clients. There are some anomalies in the data which might have to do with this aspect, such as the huge increase in active accounts and total balance amount over just a few years (see section 2). Furthermore, the assumption that the subgroup of credit card owners is representative for credit card prospects might not be valid, because ownership will be influenced by acceptance protocols, which favors clients which match the procedures, whether they are interested or not. Finally, there will always be errors which are due to the specific steps that were taken to produce the raw data. For instance, in this dataset every client starts with zero balance, which will only be true if the bank was founded on 1993, the first year that appears in the data.
The different tables were joined into a single client-oriented table. This was facilitated by the fact that there were no clients that had access to more than one account. We created different target attributes: credit card yes/no, credit card value and average number of transactions per year. All credit card owners were listed in the Cards table and received their card between 1993 and the end of 1998; there were no transactions belonging to credit card owners that received their card previously. The credit card value was derived by multiplying the average credit card withdrawal amount with the average number of credit card transactions per year. Here the assumption was made that credit card withdrawal was representative for credit card usage. The independent attributes consisted mainly of socio-demographics derived from the client's district, information on loans and totals on balance and the different transaction operations (e.g. cash withdrawal, credit in cash, transactions through other banks). The purpose of the totals was to extract frequency and monetary characteristics from the clients transaction history (min, max, average, standard deviations, counts). For credit card owners we only considered the period when they owned a card as the active period. All counts were normalized over the active period.

3.3 Mining methods

To discover interesting characteristics of target groups we used a descriptive profile analysis technique, univariate deviation detection. Assume for instance we want to compare the group of credit card owners to all clients. For a continuous attribute such as balance the procedure divides the average within the group of credit card owners by the average balance within all clients. The result is an index which illustrates the measure of interestingness for this attribute. For a nominal or ordinal attribute such as district:Prague the index is computed analogously by dividing frequencies. In the end the profiling engine checks whether the findings are statistically significant and presents the most interesting characteristics ordered by index.
The prediction models were based on the standard k-nearest neighbor algorithm (Dasarathy, 1991), including automated attribute weighting and training set condensing. Separate training and test sets were used, consisting of respectively 25% and 75% of all clients. The only free parameter (number of nearest neighbors) was set by experimentation.[2]

[2] We used the DataDetective™ visual datamining environment.

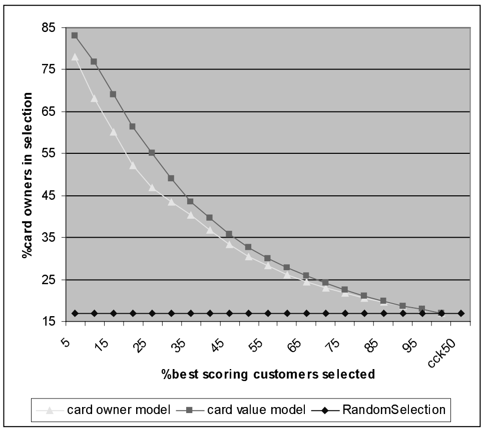


Figure 3: Response chart credit card models

4 Discovered Knowledge

The datamining results can be divided into description and prediction results.

4.1 Description results

For the cross selling business goal it is interesting to find out what makes a client a credit card customer, so we compared credit card holders to clients without credit card. Because of space restrictions we only mention a few characteristics. Credit card holders are active clients that take high cash withdrawals (factor 1.9 higher), take cash withdrawals more often (1.8) and have a higher balance (average balance 1.7, last balance 1.6). They more often have loans (1.7), but are very good at paying them back (factor 5.1; probably a side effect of a strict credit card admission policy). A number of districts were over represented, but the absolute number of clients was too small for regional effects to become statistically significant, except from the under representation of the "South Moravia" region (factor 0.81).
For the upgrading business goal we differentiated GoldCard clients from ClassicCard clients (standard card) , and JuniorCard clients from Classic Cards. GoldCard holders have a higher card value (2.3), caused by higher amounts and transaction counts per year (1.5); they collect money from another bank more often (1.5). Some findings were not significant due to the small number of GoldCard holders, but maybe interesting enough to mention: overrepresentation of North Bohemia (1.5) and North Moravia (1.4); more men (1.2). JuniorClients are of course younger (average 20 years instead of 44 years). JuniorCard Clients tend to have more cash (1.2) and credit card (1.1) transactions per year; they collect more often money from other banks (1.2). In general all other totals such as balance and transaction amounts are lower.

4.2 Prediction results

The goal in the cross selling case was to predict credit card ownership by computing a score which represents the measure of interest a client has in having a credit card. Simply predicting yes or no for 'owns credit card' does not suffice for most direct marketing purposes because it depends on the marketing tool which clients will be approached and how: top prospects get a personal visit, others a phone call, a letter, a fax or nothing.
First we used the training set to construct k-nearest neighbor models that predict a score for the 'owns credit card' attribute. If this model is used to select the top 20% prospect from the test set, the selection contains 52.2% credit card owners (see figure 3). With random selection only 17.0% credit card owners would have been found, the average percentage of credit card owners in the test set. Then we formulated the hypothesis that, though the quality of the upgrading models was not good enough to predict credit card value (remember this is average credit card withdrawal amount per year), the customer value attribute might be a better target to predict interest. The reasoning behind this was that in

reality a client owns a credit card *in a certain degree*, relative to the amount the client uses it. When the top 20% prospects are selected according to predicted customer value, the selection contains 61.4% credit card owners. So using credit card value as a target value leads to small but significant improvements in modeling credit card ownership.

4.2 Deployment and future work

The work is not finished when models have been built and evaluated. For successful datamining it is important to pay attention to the question how the discovered knowledge can be exploited and improved (see Thearling and Stein (1998) for an overview of factors for success). We focus here on exploitation of the modeling results.
For dissemination purposes, it is an advantage that we have predicted a score for credit card ownership instead of a simple yes or no answer. Instead of constructing and using all kinds of different models for different applications, the score can be entered into the core corporate database and problem specific thresholds are used in different situations. For instance, when the credit card product marketeer wants to prepare a mailing he can adjust the threshold to optimize benefits given the budget for mailing costs. Customer service operators in a call center can receive an alarm when a customer calls with a high score. If the customer informs about a travel insurance, the operator can suggest a credit card as well. In this case the threshold is probably set higher, because now it is the customer who is investing time and telephone costs. An example of a more far reaching application from the viewpoint of customer orientation would be a web site where customers themselves can check their potential for credit cards and other bank products. A particular advantage of k-nearest neighbor modeling for dynamic applications is that the models adapt automatically when the customer database changes, without the need for retraining.
To improve the value of credit card modeling it must be combined with the concept of creditworthiness. Most probably there exist a policy of strict admission rules which should be checked anyway before offering a credit card to a customer. Datamining methods can be used to model the more fuzzy concept of a 'high risk to default'. This is probably a major issue for this particular bank: for instance more than 13% of all regular loans are not fully paid back when the contract finishes; this sums up to 19% of the total loan amount.
As suggested in this article datamining could be used for customer upgrading. When more data comes available the potential of ClassicCard holders to upgrade to GoldCards could be predicted. Or alternatively, new customer value measures such as the credit card value attribute from this article could be developed to measure the value of credit card customers on a more continuous scale. Models could be built to predict the potential credit card value for a certain customer. A reward system should then be built that takes these measures into account and fosters actual and potential high value clients.

5 Conclusion

In this article we have sketched an application specific datamining scenario for promoting credit card usage, using real world transaction data. We have shown that it is feasible to develop a model to predict a useful score for credit card ownership. By sketching a datamining scenario we hope to have contributed our bit to the solution of similar problems.

References

Dasarathy, B.V. (1991). Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos, CA, 1991.

Moore, G.A. (1998). Crossing the Chasm. Capstone Publishing Limited, 1998.

Putten, P. van der (1999). Datamining in Direct Marketing Databases. In: Baets, W. (ed.). (1999). A Collection of Essays on Complexity and Management. World Scientific, Singapore, 1998.

Thearling, K., Stein, R. (eds, 1998). Keys to the Commercial Success of Data Mining: working papers. Workshop held in conjunction with KDD'98, New York., August 31, 1998.