Improving Customer Churn Prediction by Data Augmentation Using Pictorial Stimulus-Choice Data

Michel Ballings[1], Dirk Van den Poel[1], Emmanuel Verhagen[2]

[1] Faculty of Economics and Business Administration, Department of Marketing, Ghent University, Tweekerkenstraat 2, B-9000 Ghent, Belgium, http://www.crm.ugent.be, Michel.Ballings@UGent.be, Dirk.VandenPoel@UGent.be

[2] Psilogy, Leopold De Waelstraat 17a, 2000 Antwerp, Belgium, Emmanuel.Verhagen@Psilogy.com

Abstract. The purpose of this paper is to determine the added value of pictorial stimulus-choice data in customer churn prediction. Using Random Forest and 5 times 2 fold cross-validation, this study analyzes how much pictorial stimulus – choice data and survey data increase the AUC of a churn model over and above administrative, operational and complaints data. We found that pictorial-stimulus choice data significantly increases AUC of models with administrative and operational data. The practical implication of this finding is that companies should start considering mining pictorial data from social media sites (e.g. Pinterest) or use pictures instead of verbal scales in surveys, in order to augment their internal customer database. This study contributes to literature in that it is the first that assesses the added value of pictorial stimulus-choice data in predictive models. This is important because more and more social media websites are focusing on pictures.

Keywords: Customer Relationship Management, Data Augmentation, Predictive Modeling, Customer Churn, Pictorial Stimulus-Choice Data

1 Introduction

In an increasingly competitive business environment, companies have come to realize that their most valuable asset is their customer base (Athanassopoulos, 2000; Thomas, 2001). As a result, customer churn management has become the cornerstone of every customer intelligence department. From an analytical perspective, churn management consists of (1) predicting which customers are about to churn and (2) evaluating which marketing action is most effective in retaining those customers (Hung, Yen, & Wang, 2006). This study focuses on the former.

Because a customer's profitability increases over time, even small increases in retention can have substantial impact on a company's results (Gupta, Lehmann, & Stuart, 2004). Van den Poel & Larivière (2004) show that even an increase of 1%-point in retention can have a dramatic influence on contributions. Hence, companies are trying to improve customer churn prediction. In doing so, they adopt three main strategies: (1) improving data mining techniques, (2) optimizing the time window, and (3) enhancing the customer database (Baecke, & Van den Poel 2011) (Ballings & Van den Poel 2012). The latter consists in adding other data types to the internal transactional database and will be the focus of this study. Some studies also assess the value of other data types in isolation. The internal transactional database is considered the baseline for database enhancement because it contains the top predictors in extant database marketing modeling: recency, frequency and monetary value (RFM) (Baesens, et al., 2002). Since their identification (Cullinan, 1977), many studies have reported them as being the best predictors of customer behavior (e.g. Coussement, & Van den Poel, 2008a). In addition to the RFM variables, length of relationship (LOR) has also proven to be a top predictor (Van den Poel, 2003). Hence, studies that assess other data for analytical Customer Relationship Management (CRM) purposes, in addition to RFM and LOR, are considered to have a database enhancement focus.

Table 1 provides a literature review of studies that enhance the customer database in order to improve predictive models in analytical CRM.

Table 1. Literature review of data augmentation in analytical CRM

Study	Data-Variables
Steenburgh, et al., 2003	Geographical data (ZIP-codes)
Hu, & Zhong, 2008	Clickstream data
Hill, et al., 2006; Benoit, & Van den Poel, 2012	Consumer network data
Coussement, & Van den Poel, 2008b Coussement, & Van den Poel, 2009	Call center e-mails
Baecke, & Van den Poel 2011; D'Haen, et al. 2013	Commercially available survey data
Baecke, & Van den Poel, 2010	Situational variables: weather, time, sales-person variables
Thorleuchter, Van den Poel & Prinzie 2012; D'Haen, et al. 2013	Customer websites
Gilman, et al., 2004	Call center dialog transcripts
Buckinx, et al. 2007	Company survey data: loyalty
Lariviere, & Van den Poel 2004	Product features
Wong, et al., 2012	RFID

Thanks to the increasing openness of social media platforms to data extraction and even the recent rise of a social network in which pictures are the centerpiece (Pinterest) new opportunities for data augmentation are created in the realm of pictorial content. Pictures contain a massive amount of information and user actions (e.g. 'liking' a picture) could potentially reveal part of that information. Hence, this study aims at contributing to the literature by assessing the added value of pictorial stimulus choice data to customer churn prediction over and above traditional customer data. Before detailing what pictorial stimulus choice data is, the next section will first classify traditional data types.

2 Customer data sources

According to the required level of investment companies have to make, and to which customer behavior the data represents, we differentiate four customer data sources that companies can tap into for customer intelligence purposes.

As Alshawi , et al. (2011) point out, smaller or startup companies have less financial abilities to invest in information and data gathering. Because they cannot (yet) stem the costs required to store, maintain and mine huge amounts of data they focus only on gathering data about necessities (what currently needs to be delivered and invoiced). We call this administrative data because it represents customer identification and contract specifics acquired through the administrative process. The second customer data source, operational data, is the entire history of all subscriptions or contracts and operations (Larivière, & Van den Poel, 2005). Companies that do have the means to make the necessary investments linked to the extra storage, maintenance, software and skill requirements are probably more mature and bigger than companies that only use administrative data. The third customer data source is complaints data (Coussement, & Van den Poel, 2008b). This data type is related to the customer feedback process and requires a significant supplementary investment to mine given its often unstructured nature (e.g. email-mining). According to the exit-voice theory (Hirschman, 1970), complaining behavior is conceived of as one of two options when a customer is dissatisfied, next to leaving the provider, and can therefore be a valuable addition to predictive models. The fourth and final customer data source is surveys. While surveys are primarily aimed at uncovering insights (obtaining cross-sectional data for descriptive models), some companies take it one step further by using surveys as a customer data source for intelligence purposes (i.e. longitudinal data for predictive models) (Baecke, & Van den Poel 2011). This data type can be conceived of as the final category in that customer intelligence departments are adding external data from other departments, to their internal data (i.e. administrations, operations and complaints data).

In addition to these traditional sources pictures offer a new opportunity for data augmentation. As aforementioned, this is driven by the increasing openness of social media platforms and the recent rise of picture centered social networks (e.g., Pinterest). Even when used in a survey context, pictures can offer substantial advantages over traditional verbal surveys. In the following paragraphs we will provide a discussion of these advantages.

Although surveys can be an effective means for increasing the predictive accuracy of a model, it is a burden for the customer. A reason why customers don't like traditional surveys is, first of all, that many surveys are too long asking the respondent too much of his/her time. To measure a concept, e.g. satisfaction, multi-item scales are required, meaning that the length of the questionnaire increases rapidly with every concept that is measured. Second, language barriers (De Pelsmacker & Van Kenhove 2006) are often to be overcome in that the actual method of responding is inadequate for the respondent because of the incompatibility of the response method with the information that has to be conveyed; certain concepts such as emotions are often hard to be verbalized. Third, psychological or memory-based barriers (De Pelsmacker & Van Kenhove, 2006) can also pose a problem in that the respondent is induced to engage in extensive processing because the answer to a question can be hard to recall. Fourth, it is also possible that respondents are not willing to give certain answers because of social barriers (De Pelsmacker & Van Kenhove 2006) (social desirability bias, e.g. de Jong, Pieters & Fox, 2010).

There are also some problems related to traditional surveys in terms of the research itself. First, knowledge barriers (De Pelsmacker & Van Kenhove 2006) are said to exist if the respondent doesn't know the answer (e.g. when the respondent is not aware of his own attitudes or motivations). Second, when the respondent is in an analytical or rational mood and gives rational answers to emotional questions one can speak of barriers of rationality (De Pelsmacker & Van Kenhove 2006). These two limitations make traditional survey an ineffective means of acquiring certain information.

Pictures alleviate these problems in that they function as an enabling technique (De Pelsmacker & Van Kenhove 2006). Picture association is a task in which the respondent is asked to choose a picture or image out of multiple possibilities as response to a specific question. Because pictures can contain an enormous amount of information and convey it almost instantaneously to the respondent, they are a potential means for more efficient surveying. The time the customer will be solicited is shortened drastically thanks to the decrease in the time to understand the question (e.g., looking at pictures as opposed to reading multiple questions) (Dehouwer & Hermans 1994). This method can also alleviate psychological or memory-related barriers by decreasing the demand on the respondent's cognitive processing by facilitating memory retrieval (Hermans, Dehouwer & Eelen 1994). In fact, it has been empirically demonstrated that individuals' memory retrieval is more effective when the memory is based upon the same emotion as the emotional state at retrieval; this is called the mood congruency effect (Fiedler et al. 2001). As such, inducing the emotion that was experienced during encoding can enhance memory (Fiedler, et al. 2001). Given that pictures are stronger stimuli of emotion than words (Dehouwer & Hermans, 1994), showing the respondent pictures to choose from will enhance memory retrieval in a superior way relative to traditional surveys (words) if one of the pictures contains the same emotional content as the memory to be retrieved. For example, when presented with a picture of a person laughing or smiling, in contrast to simply reading the word 'joy', respondents will more easily experience the emotion 'joy' and subsequently more easily retrieve memories that are characterized by 'joy'. In addition, the other aforementioned barriers can also be alleviated by an image association technique: social barriers (pictures can be less specific than words and consequently form less of a threat for the respondent to answer), language barriers (no language is required to choose a picture), barriers of knowledge (pictures enable a respondent more to pinpoint a gut-feeling as opposed to words) and barriers of rationality (pictures are stronger stimuli of emotion than words (Dehouwer & Hermans, 1994)) (De Pelsmacker & Van Kenhove, 2006).

In sum, providing an alternative response method (pictures instead of scales) can be conceived of as being the source of multiple advantages and would therefore be an attractive means of data augmentation. To the best of our knowledge, no study has attempted this in the past.

This study aims to fill this gap in literature by evaluating the added value of pictorial stimulus-choice data in customer intelligence over and above the traditional data sources discussed above. While in traditional surveys the challenge is in providing the right answer possibilities, in their pictorial variant it comes down to showing the right pictures to choose from. 'Right' in this context means that they should induce the relevant emotional spectrum and that they should discriminate between the dependent behaviors. To accommodate this, we tried a variety of picture sets. We discuss these picture sets in the 'Variables' subsection of the following section.

3 Methodology

3.1 Data

The customers of two Belgian newspaper brands were invited to participate in a study. Both brands can be considered similar, except their geographical targeting at the province/state level. One of the two brands consisted of different editions. 25,897 emails were sent, inviting customers to click on a link to go to an online questionnaire (an incentive in the form of a prize was offered). 6,661 (25.7%) customers opened the message and of those customers 4,360 (65.9%) clicked through to the questionnaire (many dropped out after the first page). 2,605 (59.7%) of them were subsequently used in the analysis. While we did this extra survey to obtain the pictorial stimulus- choice data we also included some traditional questions such as involvement (Zaichkowsky, 1994), satisfaction (Fornell, et al., 1996), calculative commitment (Gounaris, 2005), affective commitment (Gustafsson, et al., 2005), normative commitment (Ros, et al. 1999), renewal intentions (Rossiter, 2002) and value instrumentality (Lindeman, & Verkasalo, 2005). Table 2 displays the sample characteristics (see later for cross-validation).

Table 2: Sample characteristics (5 times 2 fold cross-validation)

Across folds	Average Number of accounts	Average Relative percentage	Standard deviation
Training/Test data	1302.5		
Non-Churners	1146.94	88.023%	0.306
Churners	156.073	11.978%	0.305

We merged the survey data with internal data from which we computed variables in the aforementioned categories. As such, we were able to assess the added value of pictorial stimulus-choice data.

3.2 Time window

Depending on the length of the subscription and the promotional context, customers have to pay a certain price for their newspaper subscription. They are sent a letter reminding them that they are approaching the end of the subscription and asking them whether they want to renew their subscription, along with instructions on how to do that. Customers cannot cancel the subscription, as such churn prediction involves predicting whether the customer will or will not renew his or her subscription in the four-week period following the end of the subscription. During this period, customers keep on receiving newspapers, which they will have to pay if they renew.

In order to predictively discriminate churners from non-churners, the data needs to be analyzed according to a specific time window: independent variables are to be computed from data from the period preceding the period in which the dependent churn variable is observed. Once the model is built in the first step, both the independent and dependent period can be shifted forward in time while respecting their relative positions in order to deploy the model. Figure 1 displays the time-window.

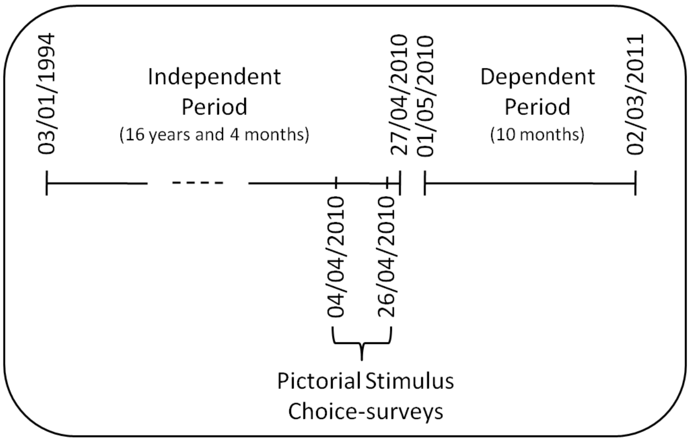

Figure 1: Time- window

3.3 Variables

3.3.1 Administrative data

Administrative data represents all information regarding agreements made between the customer and the company at the time of the purchase decision. The data is acquired at the beginning of the relationship. This entails how much, where, when and to whom the newspaper needs to be delivered. It also comprises information about the price, payment method and possible promotions.

3.3.2 Operational data

While administrative data only holds data about the current subscription level, operational data holds the entire customer history (at the subscriber's level). This means that variables are across subscriptions while they are per subscription in the administrative data and that data are acquired during the relationship, in contrast to at the beginning of the relationship.

This data also contains socio-demographic data, as opposed to customer identification data, and data about suspensions, forward interruptions, credit handling, and marketing actions, as well as response to such action (e.g., participation in games). Much of the data is not merely contractual as in the administrative data but represents a more complex commercial policy (such as credit and suspension processes) and the subsequent usage of the product (subscription) by the customer.

3.3.3 Complaints data

This data type contains information about the number and topic of complaints and the solution and answer given. Examples of the topic of complaints are: non-delivery, incomplete newspaper, delivery is too late, wrong newspaper edition. Examples of solutions are: create credit, and post delivery. Examples of answers given to customers are: force majeure, mailbox too small, mailman made a mistake, employee strike, and weather conditions. Feedback data and the Voice of Customers are acquired at customer initiated feedback moments and have been shown to be of value in predictive models (Coussement, & Van den Poel, 2008b).

3.3.4 Survey data

Survey data can be conceived of as resulting from a company-initiated feedback process. Mindset variables (e.g., purchase intentions, commitment, product recommendations), customer life style information (e.g. interests and opinions) and product evaluation data (e.g., overall satisfaction, satisfaction drivers) are examples of this type of data and are impossible to collect from internal processes. Other examples are the frequency of reading specific sections (e.g., politics, economy, culture, science, tv guide) and whether the customer reads online.

3.3.5 Pictorial stimulus- choice data

The final type of data was acquired through a choice process. This data was not available with the company so we collected it by sequentially presenting respondents with six picture sets (randomized between respondents) each containing nine pictures. The first five sets were accompanied with the following question. “Imagine a typical moment when you are reading [name of newspaper brand]. In general, how do you feel at that moment? Choose one of the nine pictures that best represents that feeling. Make your choice by clicking on the picture.”. The sets respectively contained pictures of motivational scenes, a man’s facial emotional expressions, a woman’s facial emotional expressions, geometrical forms and colors. A sixth set contained pictures of couples in different relationship stages (such as being angry, an open relationship, engagement, marriage, having children, being grandparents) preceded with the question “Which of the following pictures best describes your relationship with [name of newspaper brand]?” For all pictures and sets, dummy-variables were then created indicating whether they were chosen by the respondent or not. This resulted in 48 dichotomous variables (6 sets x 9-1 pictures). Figure 2 represents an example such as the picture sets we used (a woman's facial emotional expressions).

Figure 2: example such as the picture sets we used (the orginals were in color)

While we opted for an ad-hoc approach here (an online survey), companies could go online and mine social network data (e.g., which photo did a person 'like' out of a particular photo set?).

3.4 Estimation technique

We used Random Forest (Breiman, 2001) to create the churn model because of multiple reasons. First, literature shows that it is one of the best performing classification techniques available (Luo, et al. 2004) and is very robust and consistent (Breiman, 2001). Random Forest copes with the limited robustness and suboptimal performance (Dudoit, et al., 2002) of decision trees by building an ensemble of trees (e.g. 1000 trees) and subsequently voting for the most popular class (Breiman, 2001). Each tree is grown on a bootstrap sample with a random subset of all available predictors. Second, the method does not overfit (Breiman, 2001), which is of particular importance for this study due to the relatively large number of predictors we want to test and the small sample size. Third, variable importance measures are available for all predictors (Ishwaran, 2004). Fourth, the algorithm has reasonable computing times (Buckinx, & Van den Poel, 2005). Fifth, the procedure is easy to implement: only two parameters are to be set (number of trees and number of predictors) (Larivière, & Van den Poel, 2005; Duda, et al. 2001). We follow the recommendation of (Breiman, 2001) by using a large number of trees (1000) and the square root of the total number of variables as the number of predictors.

3.5 Model performance evaluation criteria

To evaluate the classification models' performance we used the area under the receiver operating characteristic curve (AUC or AUROC). AUC is argued to be an objective criterion for classifier performance by several authors (Coussement, & Van den Poel, 2008a; Provost, et al., 1998; Langley, 2000; De Bock, et al. 2010). The receiver operating characteristic (ROC) curve is obtained from plotting sensitivity and 1-specificity considering all possible cut-off values (Hanley, & McNeil, 1982). AUC ranges from .5, if the predictions are not better than random, to 1, if the model predicts the behavior perfectly (Baecke, & Van den Poel 2011). We use AUC instead of accuracy (Percentage of correctly classified, PCC) because AUC, in contrast to PCC, is the insensitive to the cut-off value of the 'a posteriori' probabilities (Baecke, & Van den Poel 2012; Thorleuchter & Van den Poel 2012). As such AUC is a more adequate performance measure (see Baesens, et al., 2002). AUC is defined as follows:

$$AUC = \int_0^1 \frac{TP}{(TP+FN)}\, d\frac{FP}{(FP+TN)} = \int_0^1 \frac{TP}{P}\, d\frac{FP}{N}$$

, with TP: True Positives, FN: False Negatives, FP: False Positives, TN: True Negatives, P: Positives (event), N: Negatives (non-event)

Reported AUCs are all averages over five times two-fold cross validation (Dietterich 1998; Alpaydin 1999) as explained in the following section.

3.6 5 times 2 fold cross -validation

Research has shown that the usual method to compare classification methods (t-tests to confirm significant differences in the accuracies obtained from k-fold cross validation) results in increased type-I error (Dietterich, 1998). Dietterich

(1998) and Alpaydin (1999) propose five times two-fold cross validation (5x2cv) to cope with the problem. 5x2cv randomly divides the sample (2,605 observations) in two parts (1,302 and 1,303) and repeats this process 5-times. Each part is used both as training and validation part. This process results in 10 AUC's per model (Dietterich, 1998). In order to determine whether models are significantly different in terms of AUC, we follow the recommendation of (Demšar, 2006) to use the Wilcoxon signed ranks test (Wilcoxon, 1945).

The Wilcoxon signed-ranks test (Wilcoxon, 1945) is a non-parametric test, that ranks, per data set, the differences in performance of two classifiers, while ignoring the signs, and compares the ranks for both positive and negative differences (Demšar, 2006).

4 Results

Figure 3 presents the added value per data type. The lower line is the baseline representing the data sources on the horizontal axis. The upper line represents the pictorial stimulus choice data in addition to the baseline.

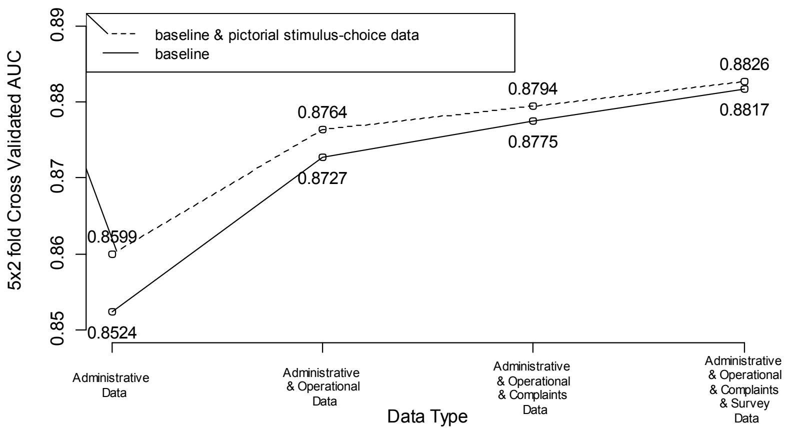

Figure 3: The added value per data type

There are three important conclusions. First, the results indicate that the addition of the operational data accounts for the biggest increase in AUC. Second pictorial stimulus-choice data improves predictive performance. Third, the performance improvement of pictorial stimulus-choice data decreases marginally when more data sources are considered.

When only administrative data is analyzed pictorials add 0.9% to the predictive performance. In the case of administrative and operational data, pictorials add 0.4%. If administrative, operational and complaints data are considered then pictorial data adds 0.2% to predictive performance. Finally, when all data are modeled (administrative, operational, complaints, and survey data) pictorial stimulus-choice data adds 0.1%.

Wilcoxon signed-ranks tests of the null hypothesis that both models (baseline versus baseline plus pictorial stimulus-choice data) perform equally well are rejected for models based on administrative data ($p < .01$), and administrative and operational data ($p < .01$). Models based on administrative, operational & complaints data are not significantly different ($p > .10$). The same conclusion can be drawn for models based on administrative, operational, complaints and survey data ($p > .10$).

In the next section we discuss the conclusions in more details. In addition we provide an outline of the practical implications.

5 Discussion and practical implications

One of the main strategies for companies to improve their predictive churn models is data augmentation. While this is an appropriate strategy for automatically recorded data (e.g., administrative data, operational data, complaints data) this is less the case if the data is the result of a conscious effort of the customer (e.g., a survey). Surveys are often considered a burden by customers and companies should be concerned with the ratio of information value by the time and effort required to capture the data. We argue that surveys using a pictorial stimulus-choice approach is a viable strategy in that pictures contain a huge amount of information, and require less time and effort from the part of the customer in the answering process. Moreover, working with pictures in contrast to answering scales can alleviate a myriad of restrictions related to cost-, time-, psychological, language, social, knowledge- and rationality- factors. In addition, social networks are the longer the more focusing on pictorial content, as testified by the recent rise of picture-centered social networks (e.g., Pinterest), which could turn out to be an important source of pictorial data for customer intelligence projects.

In this study we show that the augmentation of companies' databases with pictorial data can significantly improve customer churn prediction, even over and above a large number of traditional internal predictors. We have shown that, as expected, companies with less data to deploy (administrative and administrative & operational data) in their customer intelligence projects would benefit most from this kind of data with AUC increases of respectively 0.9% and 0.4%.

These findings have important practical implications. More and more sources of data are accessible and managers' choice-options for data-augmentation are constantly growing. Hence, managers are in need of information about the added value of these options. This study offers that information while focusing on different managerial scenarios based on the maturity of companies' data. We have two different recommendations based on these scenarios. First, we advise companies with less mature data to use pictorial stimulus choice data if they want to increase the predictive performance of their models rapidly. Second, companies with more mature data (administrative, and operational and complaints) should evaluate whether they want to ask the customer to answer a traditional survey, which is often considered to be a burden, or a pictorial survey. Admittedly, in this study traditional surveys offer a bigger increase in predictive performance (from 0.8775 to 0.8817: +0.48% versus from 0.8775 to 0.8794: +0.22%), but the time investment of the customers was also much higher for the traditional questions versus the pictorial questions (there were also much more traditional questions than pictorial questions). As aforementioned, in our opinion it is the ratio of information to time investment that is important.

6 Limitations and future research

This study intends to gauge the added value of pictures in predictive models. Since a picture contains a huge amount of information, user actions (e.g. 'liking' a picture on a social media platform) could possibly say a great deal about future behavior. Whether this is the case depends on the informational content of the picture. A limitation in this study is then that we tested only six pictorial sets. A possible direction for future research is to test more pictorial sets. The construction of optimal pictorial sets could be an interesting path to explore.

A second limitation is that we used an ad-hoc approach (an online survey). Future research could use a structural approach by mining online social networks.

A third direction for future research is to model pictorial stimulus-choice data with other dependent behavior as the criterion variable. For example, because in a customer acquisition case the available data to be incorporated in the model is far more limited than in a churn case, the added value of the image association technique could prove very valuable.

A final avenue for future research is to try different introductory questions to the pictorial sets. In this study we asked to associate a picture with the product experience, which can be conceived of as a product attribute. These attributes in turn affect the overall attitude towards the target (e.g., overall satisfaction with the product) (Guo Xiao & Tang, 2009). Subsequently, the attitude towards the target affects attitude towards behavior (e.g., commitment to staying in the relationship) (Bansal, Irving & Taylor, 2004). This attitude will in turn result in a conscious plan to engage in the focal behavior (behavioral intentions) (Solinger, van Olffen & Roe, 2008). In sum, attitudes have an impact on observed behavior, from most proximal to most distal, through behavioral intentions, attitude toward behavior (commitment), attitude toward target (satisfaction) and drivers of attitude toward target (drivers of satisfaction) (Fishbein and Ajzen 1975; Solinger, van Olffen & Roe, 2008; Eagly & Chaiden's, 1993; Bansal, Irving & Taylor, 2004; Guo Xiao &

Tang, 2009). It might prove valuable to adapt the question to gauge a more proximal attitude to behavior (e.g., Which picture best represents your commitment to stay with the newspaper?).

7 Acknowledgements

We would like to thank strategic marketing bureau [intentionally left blank] for their ideas on the topic of this paper. We would also like to thank Katrien Meert for posing for the pictures presented in this study. Finally we acknowledge the IAP research network grant No. P7/06 of the Belgian government (Belgian Science Policy).

8 References

Alpaydin, E. (1999). Combined 5 x 2cv F test for comparing supervised classification learning algorithms. Neural Computation. 11(8), 1885–1892.

Alshawi, S., Missi, F., Irani, Z. (2011). Organisational, technical and data quality factors in CRM adoption - SMEs perspective. Industrial Marketing Management. 40 (3), 376-383.

Athanassopoulos, A.D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. Journal of Business Research. 47 (3), 191–207.

Baecke, P., Van den Poel, D. (2010). Improving purchasing behavior predictions by Data Augmentation with situational variables. International Journal of Information Technology & Decision Making. 36 (3), 367-383.

Baecke, P., Van den Poel, D. (2011): Data Augmentation by Predicting Spending Pleasure Using Commercially Available External Data. Journal of Intelligent Information Systems. 36 (3), 367-383.

Baecke, P., & Van den Poel, D. (2012). Including spatial interdependence in customer acquisition models: a cross-category comparison. Expert Systems with Applications, 39(15), 12105-12113.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., Dedene, G. (2002). Bayesian neural network learning for repeat purchase modelling in direct marketing. European Journal of Operational Research. 138 (1), 191–211.

Ballings, M., & Van den Poel (2012). Customer event history for churn prediction: How long is long enough? Expert Systems with Applications, 39 (18) , 13517-13522

Bansal H. S., Irving P. G. & S. F. Taylor (2004). A Three-Component Model of Cus-tomer to Service Providers, Journal of the Academy of Marketing Science, 32, 3, 234-250.

Benoit, D. F., & Van den Poel, D. (2012). Improving custom-er retention in financial services using kinship network information. Expert Systems with Applications, 39, 13, 11435-11442.

Breiman, L. (2001). Random forests. Machine Learning. 45 (1), 5–32

Buckinx W., Verstraeten G., Van den Poel D. (2007); Predicting customer loyalty using the internal transactional database. Expert Systems with Applications. 32 (1), 125-134.

Buckinx, W., Van den Poel, D. (2005). Customer base analysis: Partial defection of behaviourally-loyal clients in a non-contractual FMCG retail setting. European Journal of Operational Research. 164 (1), 252– 268;

Coussement, K., Van den Poel, D. (2008a). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems with Applications. 34 (1), 313-327.

Coussement, K., Van den Poel, D. (2008b). Integrating the voice of customers through call center emails into a decision support system for churn prediction. Information & Management. 45 (3), 164-174.

Coussement, K., Van den Poel, D. (2009). Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers. Expert Systems with Applications. 36, 6127–6134.

Cullinan, G.J. (1977). Picking them by their Batting Averages' Recency – Frequency – Monetary Method of Controlling Circulation. Manual Release 2103, Direct Mail/Marketing Association, N.Y.

De Bock, K. W., Coussement, K., Van den Poel, D. (2010). Ensemble classification based on generalized additive models. Computational Statistical Data Analysis. 54 (6), 1535-1546.

de Jong MG., Pieters R. & Fox, JP (2010). Reducing Social Desirability Bias Through Item Randomized Response: An Application to Measure Underreported Desires, Journal of Marketing Research, 47, 1, 14-27.

De Pelsmacker P., & Van Kenhove P. (2006). Marktonderzoek, methoden en toepassingen (2de editie), Pearson Education Benelux, Prentice Hall, 2006, 515.

Dehouwer J. & Hermans D. (1994). Differences in the Affective Processing of Words and Pictures, Cognition & Emotion, 8, 1, 1-20.

Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. Journal of Machine Learning Research. 7, 1-30.

D'Haen J., Van den Poel D., & Thorleuchter D. (2013). Predicting Customer Profitability During Acquisition: Finding the Optimal Combination of Data Source and Data Mining Technique, Forthcoming in Expert Systems with Applications.

Dietterich, T.G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. Neural Comput. 10, 1895–1924.

Duda, R.O., Hart, P.E., Stork, D.G. (2001). Pattern Classification. Wiley, NY, Chapter 8.

Dudoit, S., Fridlyand, J., Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. Journal of the American Statistical Association. 97 (457), 77–87.

Eagly A, Chaiken S. (1993). The psychology of attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Fiedler K, Nickel S, Muehlfriedel T, et al. (2001). Is mood congruency an effect of genuine memory or response bias?, Journal of Experimental Social Psychology, 37, 3, 201-214.

Fishbein M. & I. Ajzen (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Reading, Mass: Addison-Wesley.

Fornell, C., Johnson, M.D., Anderson, EW., et al. (1996). The American customer satisfaction index: nature, purpose, and findings. Journal of Marketing. 60 (4), 7-18.

Gilman A., Narayanan B. & Paul S. (2004). Mining call center dialog data. In Zanasi A., Ebecken N.F.F. and Brebbia C.A. (Eds.) Data Mining V. WIT Press.

Gounaris, SP. (2005). Trust and commitment influences on customer retention: insights from business-to-business services. Journal of Business Research. 58 (2) 126-140.

Guo L., Xiao JJ. & Tang CY. (2009). Understanding the psychological process underlying customer satisfaction and retention in a relational service, Journal of Business Research, 62, 11, 1152-1159.

Gupta, S., Lehmann, D. R., Stuart, J. A. (2004). Valuing customers. Journal of Marketing Research. 41, 7–19.

Gustafsson, A., Johnson, M.D., Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. Journal of Marketing. 69 (4), 210-218.

Hanley, J.A., McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 143 (1), 29–36.

Hermans D., Dehouwer J. & Eelen P. (1994). The Affective Priming Effect, Automatic Activation of Evaluative Information in Memory, Cognition & Emotion, 8, 6, 515-533.

Hill, S., Provost, F., Volinsky, C. (2006). Network-based marketing: Identifying likely adopters via consumer networks. Statistical Science. 21, 256–276.

Hirschman, A.O. (1970). Exit. Voice, and Loyalty—Responses to Decline in Firms. Organizations, and States. Cambridge, Harvard University Press.

Hu, J., Zhong, N. (2008). Web farming with clickstream. International Journal of Information Technology & Decision Making. 7, 291–308.

Hung, S.-Y., Yen, D.C., Wang, H.-Y. (2006). Applying data mining to telecom churn management. Expert Systems with Applications. 31, 515–524.

Ishwaran, H., Blackstone, E. H., Pothier, C. E., Lauer, M. S. (2004). Relative risk forests for exercise heart rate recovery as a predictor of mortality. Journal of the American Statistical Association. 99 (467), 591–600.

Langley, P. (2000). Crafting papers on machine learning. In: Langley, P. (Ed.), Proc. of 17th International Conference on Machine Learning. ICML-2000. Stanford University, Stanford.

Lariviere B., Van den Poel D. (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. Expert Systems with Applications. 27 (2), 277-285.

Larivière, B., Van den Poel, D. (2005). Predicting customer retention and profitability by using random forests and regression forests techniques. Expert Systems with Applications. 29 (2), 472-484.

Lindeman, M., Verkasalo, M. (2005). Measuring values with the short Schwartz's value survey. Journal of Personality Assessment. 85 (2), 170-178.

Luo, T., Kramer, K., Goldgof, D.B., Hall, L.O., Samson, S., Remsen, A., et al. (2004). Recognizing plankton images from the shadow image particle profiling evaluation recorder. IEEE Transactions on Systems, Man and Cybernetics Part B-Cybernetics, 34 (4), 1753–1762.

Provost, F., Fawcett, T., Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In: J., Shavlik (Ed.), Proc. of 15th International Conference on Machine Learning. ICML-1998. Morgan Kaufman, San Francisco, CA.

Ros, M., Schwartz, S.H., Surkiss, S. (1999). Basic individual values, work values, and the meaning of work. Applied Psychology-An International Review. 48 (1), 49-71.

Rossiter, JR. (2002). The C-OAR-SE procedure for scale development in marketing. International Journal of Research in Marketing. 19 (4), 305-335.

Solinger, O. N., van Olffen W. & R. A. Roe (2008). Beyond the three-component model of organizational commitment, The Journal of applied psychology, 93, 1, 70-83.

Steenburgh, T. J., Ainsle, A., Engbretson, P. H. (2003). Massively categorical variables, revealing the information in ZIP codes. Marketing Science. 22, 40–57.

Thomas, J.S. (2001). A methodology for linking customer acquisition to customer retention. Journal of Marketing Research. 38 (2), 262–268.

Thorleuchter D., Van den Poel D. & Prinzie A. (2012). Analyzing existing customers' websites to improve the customer acquisition process as well as the profitability prediction in B-to-B marketing. Expert Systems with Applications. 39 (3), 2597-2605.

Thorleuchter, D., & Van den Poel, D. (2012). Predicting e-commerce company success by mining the text of its publicly-accessible website. Expert Systems with Applications, 39(17), 13026-13034.

Van den Poel, D. (2003). Predicting mail-order repeat buying: Which variables matter? Tijdschrift Voor Economie En Management. 48 (3), 371–403.

Van den Poel, D., Larivière, B. (2004). Customer attrition analysis for financial services using proportional hazard models. European Journal of Operational Research. 157, 196–217.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics. 1, 80–83.

Wong W. K., Leung S. Y. S., Guo Z. X., Zeng XH., Mok, P. Y. (2012). Intelligent product cross-selling system with radio frequency identification technology for retailing. International Journal of Production Economics. 135(1), 308-319.

Zaichkowsky, JL. (1994). The Personal Involvement Inventory- Reduction, Revision, and Application to Advertising. Journal of Advertising. 23 (4), 59-70.