Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2024) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

# Enhancement of a costumer targeting model by social network analysis and data mining

Sebastián A. Ríos[a], Gonzalo Hermosilla[a]

[a]*Business Intelligence Research Center, Industrial Engineering Department, University of Chile, Santiago, Chile*

**Abstract**

Costumer targeting is becoming critical for successful advertising and client retention. In this case we studied the performance of a targeting model in terms of adoption and profitability. The base model chooses a set of clients to offer them mobile products through a telephone campaign. Then, we proposed an enhanced model that incorporates social attributes extracted from the mobile phone network of each client. Additionally, we studied the influence that friend's prior adoption might have on their own adoption.

It was discovered that the incorporation of social attributes does not improve the power of prediction of targeting models, but instead, what improves it is the combination of both models (base and enhanced model). Hence, to achieve effective targeting models it is necessary to combine models based on different classification techniques in order to increase the power of prediction. In this case, data aggregation in the training set and the combination of both models allow us to increase in 8% the model performance on average, reaching a level of hits of 89%.



*Keywords:* Targeting; Telecommunications Industry; Big Data; Social Network Analysis; Support Vector Machine

## 1. Introduction

Over the years, companies have implemented a set of processes and systems which allow them to raise many business strategies. These strategies are focused on building a profitable and long-term relationship with specific customers. This is known as Customer Relationship Management (CRM)[9]. Thanks to the development of Internet and new technologies, CRM has focused on business intelligence as a tool for acquiring and retaining customers. Hence, CRM allow companies to maximize the value that each customer brings to the organization[11].

* Corresponding author. Tel.: +56-2-2978-0545
*E-mail address:* srios@dii.uchile.cl

1877-0509 

Generally, there are three aspects needed to effectively increase the customer value: up-selling, cross-selling and retention[14]. Up-selling is to sell the customer similar products bought earlier but with higher value; cross-selling refers to sell products that the customer has never purchased, which means a new type of product; retention corresponds to the effort made towards keeping customers within the company.

In Telecoms industry, CRM appears as one of the fundamental tools to differentiate between companies. As mentioned in [13], it is generally more expensive to get a new customer than to retain old ones, so the efforts must focus on retaining and increasing their value. Then, CRM aims to select customers who could potentially increase their value based on Business Intelligence and Data Mining techniques. Using classification models such as Decision Trees[15], Neural Networks[10] and Support Vector Machines[3], it is possible to identify the most likely or more likely customers to adopt a product already known through up-selling (ex. increase the value of your mobile phone plan) or a new one through cross-selling (eg. buy a data prepaid card).

A new approach is the social targeting. It is focused on considering the social value of each customer based on the relationship with the members of its own social network [5]. Since the mobile network of a customer can be represented as a social network[6], the use of Social Network Analysis (SNA) and Data Mining techniques makes it possible to extract information that can be used to create "social attributes". Those attributes can complement classification models already known for targeting customers.

The main idea behind incorporating social value is that a customer cannot qualify for a campaign considering only its individual value. In addition, the customer can have many friends into its network who already adopted the product of the campaign. Therefore, those friends could create a "social pressure" over the customer, increasing its probability of buying the product offered by the campaign.

Our main objective is to improve the prediction of a targeting model for the adoption of products through the inclusion of social attributes extracted from a social network of mobile phones, using SNA and Data Mining. We test social targeting models using three different classification techniques. We compare the results between them and against the results obtained with a model based just in commercial attributes.

## 2. State of the art

This research combines two worlds: on one hand, the traditional or classical approach of targeting (which falls into the classification area) and on the other hand we have the social network analysis world. Of course, most companies already use any classification algorithm (random forests, support vector machines, multilayer perceptron, etc.) for targeting purposes; therefore, we are not going deep in this aspect; however, how companies have been using social attributes for targeting purposes is more novel and we will explain more in the following.

### *2.1. Customer Social Targeting*

Direct marketing campaigns are developed based on customer targeting. They aim to select those customers that are potentially more profitable and focus only on them. However, this strategy provides a basic but very important limitation: treat each customer as if they take decisions independently from other clients [5]. In the real world, a person's decision to buy or purchase a product is usually strongly influenced by friends, acquaintances, relatives, business associates, etc. This influence between the client and its social network is the main idea behind the concept of social targeting, which depends on the attributes that can be derived from the configuration of the network, being represented through the social value that each subscriber owns.

Marketing campaigns based on word of mouth can be much more effective in terms of cost than traditional campaigns, since customers do much of the promotional effort.

In [7] are described three complementary strategies for targeting of customers based on SNA can be focused to achieve marketing objectives:

- Explicit Recommendation: Customers are supporters of the product or service, recommending it to their friends or acquaintances. An example of this is the success of a book that was given free of charge to 10,000 influential readers (booksellers) in order to stimulate the sale of the same.

- Implicit Recommendation: Even if people do not talk about the product, they can implicitly suggest through their actions, especially through their own adoption of the product. An example of this may be the efforts made by companies to show famous athletes using its brands, influencing their followers to also use them. Companies try to replicate this effect on small groups of people, trying to persuade the leader member to adopt the product.
- Network Targeting: The firms try to focus the marketing campaign on the residents of the social network of customers who previously adopted the product. Even if there was no recommendation or endorsement of it. An example of this was the strategy used by Hotmail where every mail sent by any user had a footer inviting the recipient of the message to take the free product.

There are several ways to address the problem of social targeting of clients, whose main challenge is to find the best way to incorporate the social value of each customer in the process of identifying the target segment. One of them focuses on the idea of combining the user attributes with social attributes that can be drawn from the structure of its social network. Regarding those attributes it is possible to build models with more predictive power in terms of identifying future adopters of a product [1, 2]. This theory is tested in a social network of instant messaging, for the adoption of a product and for the common interest on an advertisement.

Another way to assess the importance of the social network of each client in their decisions corresponds to the notion of *adopting social value*. This value is derived from the behavior of individuals within your network, since the decision to adopt a product is generally influenced by friends, colleagues, etc.

There are two different approaches to calculate the value of social adoption of a client. On [5] it is proposed a method which states that this value derives from the influence that the customer may have on other customers in the future. In turn, in [2] it is proposed the existence of an influence on an individual from their friends who previously adopted the product in question (predecessors), which creates a social pressure that can lead to buy the same product.

Both ways warn that if the social value is not taken into account in a customer targeting process, the model can reach suboptimal results in predicting future adopters. This means that a client cannot be selected for a direct marketing campaign since its individual value turns out to be less than the cost of it. However, if the social value is considered, it may make the cost fall and getting the client included in the selection. Hence, this concept is very important to the targeting of customers.

Our work considers both approaches: attributes of social value and the value of social adoption.

## 3. About the Data

We have demographic and commercial behavior data provided by a mobile phone company on a monthly basis. Demographic data includes personal customer information, such as age, gender, place of residence and socioeconomic status. It has been anonymized for the purposes of the study. The commercial behavior data includes how long the customer has been related to the company, amount of calling minutes used, number of text messages sent, the use of the internet service, and the billing amount. Each client record has 177 variables. The company has a costumer base of nearly 9 million subscribers. The data collected covers the months of July, August and September of 2013 and applies only to those customers who have some kind of mobile phone plan contracted with the company, reaching a number close to 2.7 million customers per month.

In addition, we have information about the three marketing campaigns undertaken by the company during this period summarized in Tables 1,2, and 3, each targeting product to a specific group of clients: up-selling of contract, i.e. an improvement in the current contract; up-selling contract with controlled account; and cross-selling aimed at selling a “bag of bytes” for mobile navigation. These campaigns were carried out by a call center channel to reach the target customer. Information includes if the costumer was contacted and accepted the offer.

Additionally, we have two databases that store social information about customers. One shows detail about the relationships that are built into the social network of mobile phones, represented through a social graph. It has about 5 million customers and over 125 million links and specifies the importance level of the relationship. From it, 17 social attributes that characterize each customer are extracted and stored in the second database. This data is related to the number of client’s contacts (strong and weak) and communities it belongs to.

| Product | In Campaign | Contacted | Positives | % Positives |
|---|---|---|---|---|
| Contract | 118,659 | 36,310 | 4,296 | 3.6% |
| Controlled account | 103,977 | 38,802 | 3,698 | 3.6% |
| Mobile navigation | 54,817 | 14,909 | 1,150 | 2.1% |

Table 1. Statistics of July's campaigns

| Product | In Campaign | Contacted | Positives | % Positives |
|---|---|---|---|---|
| Contract | 174,738 | 16,692 | 4,059 | 2.3% |
| Controlled account | 104,001 | 11,368 | 7,353 | 7.1% |
| Mobile navigation | 48,824 | 3,500 | 1,284 | 2.6% |

Table 2. Statistics of August's campaigns

| Product | In Campaign | Contacted | Positives | % Positives |
|---|---|---|---|---|
| Contract | 181,068 | 28,402 | 5,954 | 3.3% |
| Controlled account | 112,703 | 29,260 | 7,104 | 6.3% |
| Mobile navigation | 52,014 | 9,049 | 1,501 | 2.9% |

Table 3. Statistics of September's campaigns

| Product | Corr > 0.9 | Corr > 0.7 |
|---|---|---|
| Contract | 106 | 68 |
| Controlled account | 110 | 68 |
| Mobile navigation | 103 | 69 |

Table 4. Attributes selection considering correlations over a 0.9 and 0.7 respectively using data from July.

### 3.1. *Integration and Reduction of data*

In order to build the definitive database that will be used in training classification models, we had to integrate the three databases: socio-demographic and commercial customer behavior variables; the attributes detailing customers response to the offer made by the company; and data including social attributes.

The first step is to add the databases that have attributes of customers, i.e. gather socio-demographic, commercial and social variables to form a database of 119 variables. Following this, we added the newly built variable with the response of each customer to the offer. This attribute will be the target in the classification model. The model must learn to find patterns that allow to identify customers likely to respond positively.

Additionally, to avoid redundancies existing in the integrated database, we performed a selection of attributes based on an analysis of correlation between variables. At first, this was done only for the data of July. From the above, two new configurations of data are built: one where the variables have a coefficient of correlation greater than 0.9 and another one in which the attributes have a coefficient higher than 0.7. In both cases, from a pair of correlated variables, the variable representing a total is kept and the variable which has partial information is deleted (e.g. `minutos_totales` would remain and `minutos_compania1` would be eliminated). Table 4 details the amount of selected attributes that will be used in the classification models.

### 3.2. *Data Balance*

Since in this problem the number of clients who accept the offer for a product is very small compared with the number of customers who do not, we have a clear imbalance of classes which is showed in Table 5; all months have a very similar imbalance. To fix this problem we used the SMOTE oversampling technique and also a randomly

| Product | Positive class | Negative class |
|---|---|---|
| Contract | 3.6% | 96.4% |
| Controlled account | 3.6% | 96.4% |
| Mobile navigation | 2.1% | 97.9% |

Table 5. July imbalanced classes to be corrected by SMOTE

| Phone1 | Phone2 | Call12 | Call21 | SMS12 | SMS21 | Dur.Call12 | Dur.Call21 |
|---|---|---|---|---|---|---|---|
| 111 | 112 | 1 | 1 | 0 | 0 | 12 | 190 |
| 113 | 114 | 0 | 0 | 1 | 0 | 0 | 0 |
| 114 | 116 | 1 | 0 | 0 | 0 | 68 | 0 |
| 115 | 116 | 2 | 1 | 0 | 1 | 79 | 23 |
| ... | ... | ... | ... | ... | ... | ... | ... |

Table 6. Details of transactions using bidirectional links

subsampling in order to reduce the number of instances of negative class. We want to compare the results with both techniques.

In a first case, we used proportions of 50% for each class and in a second case, 30% for the positive class and the remaining 70% for negative class. As before, in a preliminary way, this was done only for the month of July. With SMOTE technique it was necessary to make a 100% over-sampling to Contract and Controlled Account, and 300% for Mobile navigation, plus a 200% under-sampling for the case of 50%–50% for both first products and 150% for Mobile navigation. For the case of 30%–70% the same over-sampling was used and a 400% under-sampling for both Contract and Controlled Account and 300% for Mobile navigation. The subsampling was necessary so as to not create a training set that were built too large from only synthetic instances.

## 4. Extraction of social attributes

### *4.1. Creating the social network*

As we explained before, data with telephone transactions is stored in databases called CDRs, which save a number of details of phone calls (type 1) or text messages (type 2).

As a first step, we added the details of all transactions generated in the studied period, summarizing all the interactions that are generated from sender to receiver. At this stage, there are still two rows describing the relationship between two users, each one representing this relationship in one of two possible directions: two opposite unidirectional links.

It is essential that the interaction between sender and receiver is described in a single row. So we combined the data to generate a single bidirectional link between two users. Table 6 shows how the database looks like before it is transformed into a graph. We remark that the treatments and transformations that are under these databases replicate the procedures performed by the telecommunications company. In order to generate comparable graphs where the difference in the results of the experiments are only produced by the application of targeting models and not by differences in these settings.

To achieve the extraction of attributes from the network it is necessary to build a graph containing a weight and directionality for each link. It will be calculated from the data stored in the complete database corresponding to a full month. The weight represents the importance of the relationship between two users and it is calculated based on the total amount of calls and text messages of the month and can present values greater than 1. The directionality reflects what user is responsible for the weight of the relationship. It will be calculated based on the share of calls and messages from sender to receiver of all interactions that are generated between them throughout the month. It may

| Phone1 | Phone2 | Weight | Direction |
|---|---|---|---|
| 111 | 112 | 0.47 | 0.9 |
| 111 | 115 | 0.16 | 0.38 |
| 113 | 114 | 1.26 | 0.24 |
| 113 | 116 | 1.87 | 0.65 |
| 115 | 116 | 0.82 | 0.73 |
| ... | ... | ... | ... |

Table 7. Example of a social graph describing the interaction between mobile phones

adopt values between 0 and 1. In Table 7 it can be seen how the database representing the social network of mobile phones graph is constructed.

$$Weight = \frac{\sum_{i=1}^{MonthDays} call_i + \sum_{i=1}^{MonthDays} sms_i}{4} \tag{1}$$

$$Direction = \frac{\sum_{i=1}^{MonthDays} call12_i + \sum_{i=1}^{MonthDays} sms12_i}{\sum_{i=1}^{MonthDays} call_i + \sum_{i=1}^{MonthDays} sms_i} \tag{2}$$

Replicating the configuration adopted by the telecommunications company, the relevant connections to the graph used in this study are those that have a weight greater than or equal to 1, i.e. customer relationships involving 4 or more interactions per month, either through phone calls, text messages or multimedia messages. Given this, it was decided to filter the graph and discard all links that had a weight less than 1, being structured from approximately 5.2 million nodes and 13.5 million edges. This implied a significant reduction in its structure, which would allow an easier management. Anyway, there are other ways to generate more elaborated thresholds to filter graphs, as described in [16].

### 4.2. Extraction of attributes

The social attributes of each client were removed from the social network for mobile phones described in the previous section. The variables related to the number of customer contacts were obtained by counting links that each possessed, differentiating contacts between strong and weak depending on the weight. Strong links are those with a weight greater than one and weak links have a lower weight. In this work, attributes related to the communities the customer may belong to were extracted through the discovery of cliques, which correspond to the maximum subgraph composed of at least three nodes where all nodes are connected to each other [4].

The best known method to find cliques is proposed by [12] and called percolation of cliques. It is based on the probability of internal edges of a community forming cliques. On the one hand, it is likely to form cliques if communities are dense in edges internally. Secondly, it is probably to not form cliques between communities if there are not many edges between them. K-clique is the term used to refer to a subgraph of k elements forming a clique. Then, it is possible to determine which two k-clique are adjacent if they share k-1 nodes. Thus, it is defined that a chain of k-cliques corresponds to the junction of adjacent elements, and then it is defined a community of k-cliques as the subgraph with the largest number of them connected. Then, to find these k-cliques within the network it is necessary to design a heuristic, which will start building the communities from the above definitions.

## 5. Experiments

Targeting models were applied on three products that the telecommunications company offers: up-selling contract, up-selling controlled account and cross-selling navigation on mobile account. We used three classification techniques with different approaches: Random Forest, Multilayer Perceptron and Support Vector Machines. The execution of these models was framed in many experiments designed to compare the results of the different configurations of the models and to find the one with the greatest predictive power of product adoption.

| Experiment | Training Data | Data to Predict |
|---|---|---|
| Experiment 1 | July data | August |
| Experiment 2 | August data | September |
| Experiment 3 | July data | September |
| Experiment 4 | July and August data | September |

Table 8. Experiments done to evaluate the predictive power of models

Experiments were designed to change the structure of the training set that was going to be used as input for classification algorithms (the basis for targeting models). Table 8 details each experiment that was carried out.

The main objective was to identify which configuration would be the best to find the greatest amount of customers likely to adopt a product offered by the telecommunications company, in order to attract those users using a lower amount of financial resources.

## 6. Social Targeting of clients

With the design of the experiments ready and databases completed, we proceeded to apply the different classification techniques: Random Forest Techniques, Multilayer Perceptron Neural Network and Support Vector Machines. They were executed on the training data constructed for the three products that are part of the research (Contract, Controlled Account and Mobile navigation).

For these three products, classification techniques used generated a list with two scores for each client on the testing data. These scores represent the probability that each customer has to belong to one of two classes. So, every score is between 0 and 1. The sum of these results is exactly 1, where the positive class corresponds to customers who accepted the offer and ended up adopting the product offered, while the negative class is for customers who didn't provide an affirmative answer regarding the offer.

Once the process of getting the scores was finished, we built a customer ranking score based on the score that each had for the positive class. In this particular case, the ranking was structured as a list of customers ordered from highest to lowest probability of belonging to the positive class.

After the construction of the ranking, cuts were made on top of it (higher scores) considering different proportions of clients. The idea was to analyze how the predictability of classification models evolves as the number of customers considered increases. Specifically, cuts were generated in 10%, 20%, 30% and 40% higher in the ranking. So, for each product we had six different rankings depending on the classification technique and the balance of classes in the training data. This allowed us to compare and analyze which combination was better to predict products adoption for clients.

As a test, we carried out a process of conventional classification (class approach, not scores) on July dataset to assess whether the selection of attributes such as balancing classes would allow better results than when these strategies had not been used. For this, the database of July was separated into a training set and a test set in proportions of 1/3 and 2/3 respectively, as proposed in [8]. Based on the results for the month of July, we decided which of the alternatives for modeling August and September would be used. According to this, the four experiments described were designed.

The first step was to calculate the value of social adoption of a customer as the amount of friends he had in the month of September. It was necessary to count how many users were linked to the client with a weight higher than 1. After the number of friends per customer was known, we identified how many of them had previously adopted the product. This was achieved by analyzing the graph and counting the number of friends who bought the product within a period of two months immediately prior to the studied months. In this case it corresponded to July and August. With this data we calculated the score of social adoption, which is the number of friends that adopted the product in the period divided by the total number of friends of the client.

The value of social adoption was combined with the individual score calculated for September by adding both values and normalizing them (between 0 and 1). This procedure generated a new ranking of customers. The same

| Configuration | Recall DT | Recall NN | Recall SVM |
|---|---|---|---|
| Under(50/50) | 63.6% | 57.8% | 34.1% |
| Under(50/50)_FS_0.9 | 63.3% | 57.3% | 32.4% |
| Under(50/50)_FS_0.7 | 62.1% | 50.6% | 29.5% |
| SMOTE(50/50) | 18.9% | 39.9% | 26.2% |
| SMOTE(50/50)_FS_0.9 | 17.8% | 37.1% | 24.7% |
| SMOTE(50/50)_FS_0.7 | 16.4% | 36.6% | 23.8% |
| SMOTE(30/70)_FS_0.9 | 0.69% | 24.5% | 0% |
| SMOTE(30/70)_FS_0.7 | 0% | 14.1% | 0% |

Table 9. Results of classification done on July data for Contracts

assessment methodology wsa used, just as before, based on the ability to identify new clients adopting the product and the proportion of accumulated successes.

## 7. Results

### *7.1. Client targeting*

Considering the three classification techniques and the two strategies for balance, we built six rankings for each product in each month in the study. To find out which one represents the model with the greatest predictive power, we calculated recall, precision and f-measure. They were calculated for groups of related customers at four cuts in the ranking (10% to 40% from the total).

To summarize we present results of one experiment on Table 9, since we obtain similar results on the other experiments. Here we see that the strategy of selecting a set of attributes eliminating all variables that have high correlation does not produce improved results. So, we chose to keep all the variables for modeling the other months. As it is observed, the best results are obtained working with a balance of classes by 50%-50%. Based on these results, it was decided to continue with the experiments without using attribute selection and performing a balancing way to match the weight of both classes (both over-sampling and under-sampling).

After we defined the configuration of the data to be used, we continued with the experiments 1 and 2. The first results are depicted in Figure 1; due to space purposes we can not show all developed tables and figures, however, we will explain our conclusions in the following. Through bar charts we show the proportion of customers who adopted the product identified only by the base model[1] (blue portion), customers found only in social models (dark blue portion) and customers detected by both models (light blue portion). This process was repeated for the three classification techniques.

Results indicate that the product Contract (figure 1) —from the model proposed using Random Forest (DT)— has in almost all cuts of the ranking a higher level of precision compared to the base model. This can be seen more clearly in the prediction for the month of August. The total number of customers of the positive class identified is higher and the amount of "new" customers who are found by the proposed method (and not by the base model) is also higher (independent of class balancing method used). However, Multilayer Perceptron Neural Network (MPNN) and Support Vector Machines (SVM) classifiers show a similar or lower level than the method without social attributes. This is deducted from fewer customers identified and fewer new customers tried to detect.

On the other hand, regarding the product Controlled account , it is possible to note that for both months (August and September), the predictability of the social model is very similar for the three classification techniques used. The results show that regardless of the method of balancing classes, in the month of August the social model can identify a greater number of new customers compared to the base model, especially when cuts 10% and 20% are analyzed. However, for September, even if we increase the cut proportion for the ranking to detect a greater number of

[1] The base model corresponds to the model with socio-demographic and behavioral variables, excluding social attributes.

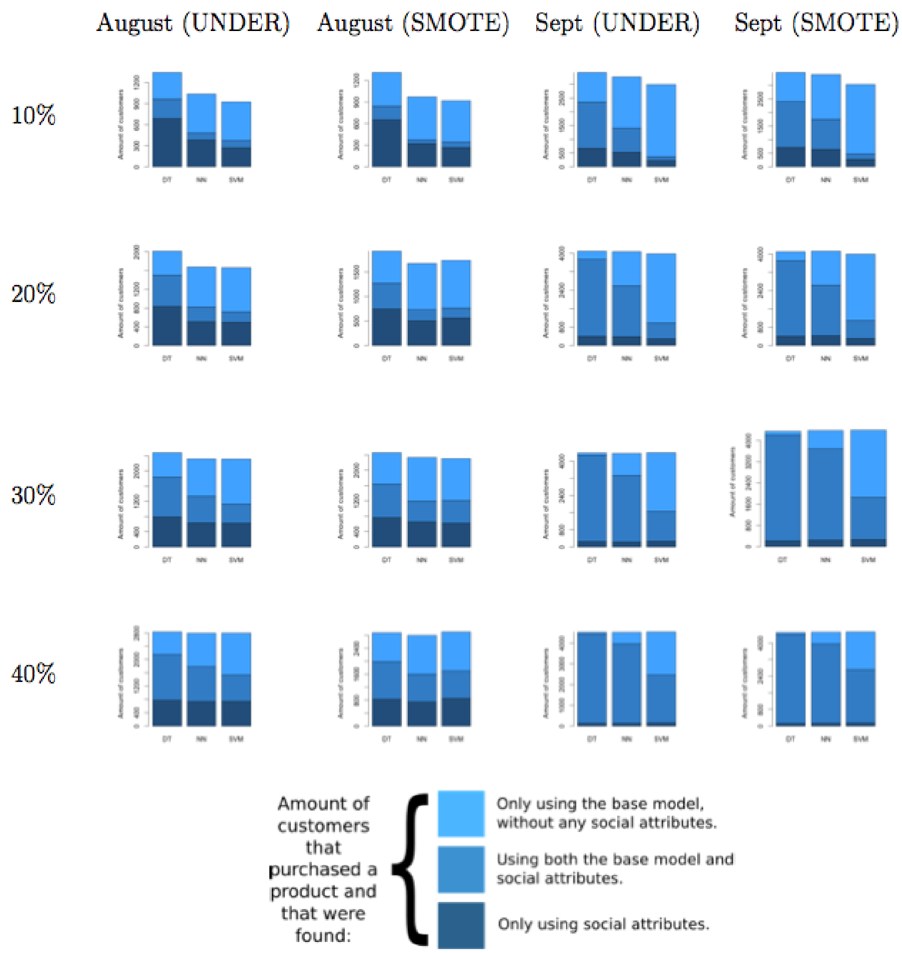


Fig. 1. Results for experiments 1 and 2 for "Contract" product.

customers, the amount of new customers detected decreases. Among the techniques used, the best performance with balance under-sampling are MPNN for the month of August and DT for September. For over-sampling are SVM and DT respectively.

Finally, in the case of Mobile Navigation, the results for the month of August are generally below the results shown by the base model for both: the balanced case with undersampling and with over-sampling. The exception is for the 10% cut where the proposed model can detect a significant number of new customers. This fraction decreases as the proportion of cut increases. Opposite to this, for the prediction of September, the Random Forest model achieved very positive results, finding more customers than the base model and identifying a large number of new customers not previously detected. These results are repeated for the two methods of balancing classes.

Another idea, that would be useful to evaluate the predictive capacity of the proposed models, is related to the use of the so called recall measure. It is possible to analyze how it will increase the number of hits as the cut takes a higher proportion of total customers. This form of evaluation is known as the cumulative curve hits.

### *7.2. Client's social targeting*

This section presents the results obtained after including the value of social adoption into the rankings for customers calculated in the previous experiments. It is compared against the results generated in experiment 4.

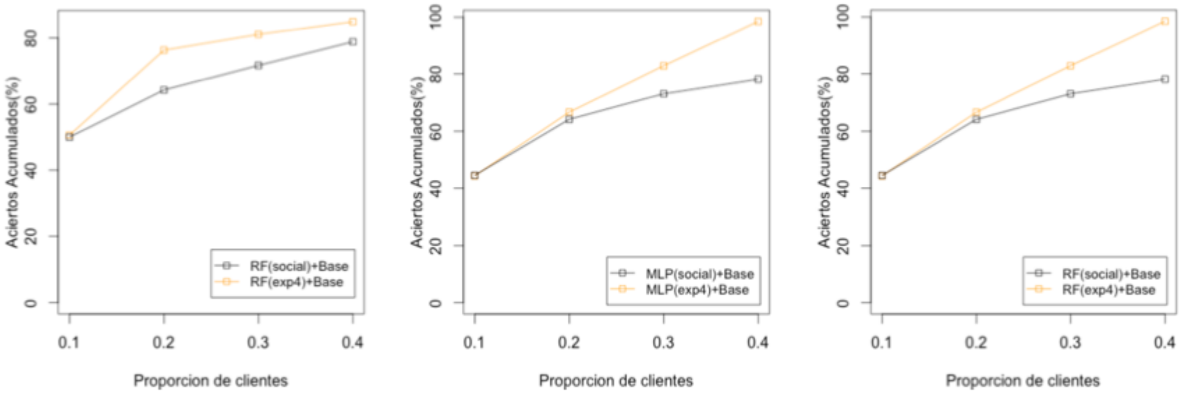


Fig. 2. Comparison of cummulative results of experiment 4 and the inclusion of social adoption value for Contract.

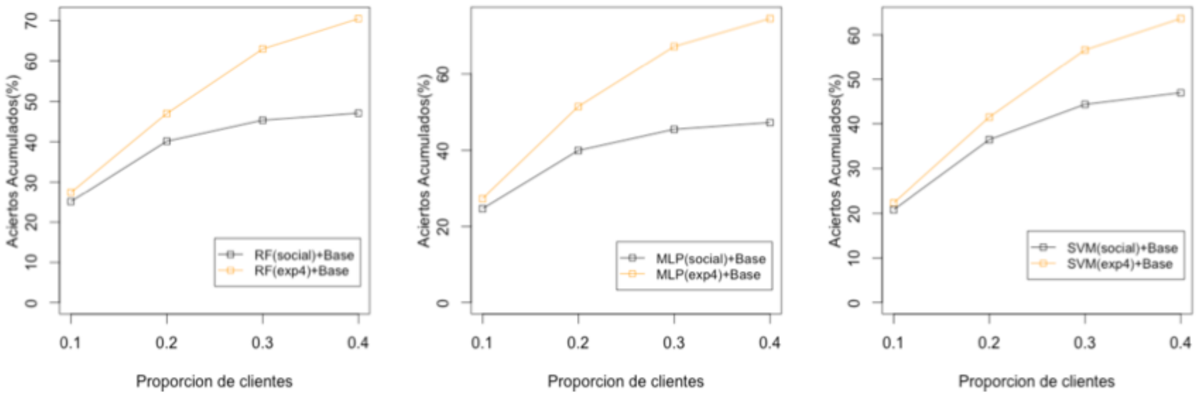


Fig. 3. Comparison of cummulative results of experiment 4 and the inclusion of social adoption value for Controlled Account.

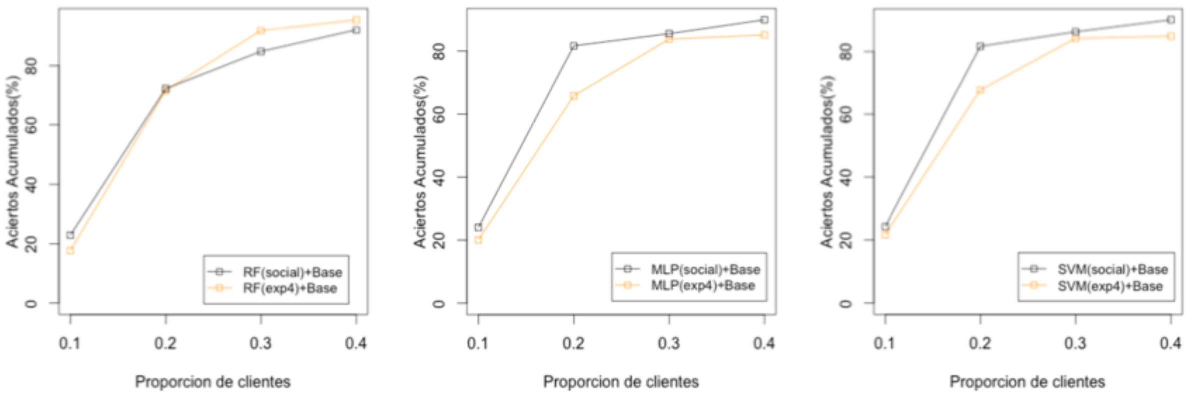


Fig. 4. Comparison of cummulative results of experiment 4 and the inclusion of social adoption value for Mobile Navigation.

It is observed that for Contract and Controlled Account products (figures 2 and 3), the inclusion of the value of social adoption in the ranking of clients (black curve) generates a fall in the amount of accumulated hits along the cuts. This is independent of the technique used for classification. For Contract, the percentage of hits falls around 15% on average for the three classification techniques, where the strongest decrease occurred with SVM (-20.27%). In the case of Controlled account, the average drop in the predictive power reaches 34%, where the sharpest decline was using SVM (-51.4%).

For Mobile Navigation (figure 4, the results show that in this case the inclusion of the value of social adoption matches and even improves the results obtained in experiment 4. Specifically, when Random Forest is used, the results from both models are almost identical, showing a slight advantage for experiment 4. However, with MLP and SVM techniques, the advantage is given by the incorporation of the value of social customer adoption, which allows models to identify about 5% of more customers.

Note that for all three products and three classification techniques, results obtained after cutting the highest 10% are the same for both experiment 4 and the model with modified ranking clients using the value of social adoption. But as cuts incorporate a larger proportion of customers, the difference between the results of both models gets increased.

## 8. Conclusions

The first results obtained show improvements in identifying customers who adopt the products offered by a marketing campaign, although not in the way that was initially expected. At the beginning it was thought that the model of targeting would achieve better results for the mere fact of incorporating information about social attributes. The first two experiments conducted showed that this was not achieved because the power of prediction of the proposed social model was similar or even worse than the base model, a phenomenon that was repeated for all three products under study, independent of the classification technique used. However, when analyzing the results from the point of view of *what* and not *how many* clients were being identified, the first key discovery of this work was achieved: the targeting social model and the base model are able to select customers with different profiles. This important conclusion tells us that by combining the hits predicted by both targeting models, we can achieve a significant increase on the percentage of customers who are identified as adopters.

Whereas the variation in the design of a targeting model allows us to identify different customer profiles, it could be the case that under the use of several models, with each model pointing to a different profile, all clients prone to adopt any mobile phone product are identified. It would mean an important improvement for direct marketing campaigns as they would approach to an accuracy close to 100% accuracy. This fact means significant savings for companies. To achieve this efficiently, the intersection between the sets of customers found by the different models should be as small as possible.

A second key finding of this work is done by analyzing the results of experiments: setting the training set to be the combination of customer data from the two previous months, the results show a new improvement in the predictive power of the model. Again, this is achieved for all three products studied, regardless the classification technique used.

It is important to note that the datasets are combined to form the training set corresponding to consecutive periods. This requirement allows the model to identify the hidden trending pattern in the attributes associated with the client. Otherwise, it could happen what was observed in experiment 3, where the training set built from a dataset two months old ($t - 2$) generates a not so significant improvement as was the case of experiment 4.

Regarding the classification techniques used, while Random Forest is the technique that generally performs best —in 61% of the cases studied it is the one with the greatest power of prediction— this behavior prevails only when the achievements of the social model are considered and not combined with the based model. The evidence of this lays in experiments 1 and 2, where Random Forest algorithm proved to stand out above the others in the three products studied. However, when the combination of the hits is considered, the four experiments achieve better results with MLP and SVM techniques. This phenomenon shows that while these algorithms fail to identify the greatest number of customers, they find more clients for different profiles than Random Forest and the base model.

This reinforces the idea that it is necessary to work with several complementary models focusing on customers with different profiles to achieve a more effective and efficient prediction. Combining targeting models designed from different classification algorithms, it is possible to build an aggregate model that raises its prediction capability to levels close to 100%, allowing companies to increase the performance of their marketing and business plans.

As future work, we propose to test other classification techniques, incorporating more data (more months) and testing new approaches to include the value of social adoption in rankings.

## References


[1] Bagherjeiran, A., Parekh, R., 2008. Combining behavioral and social network data for online advertising, in: Data Mining Workshops, 2008. ICDMW'08. IEEE International Conference on, IEEE. pp. 837–846.

[2] Bhatt, R., Chaoji, V., Parekh, R., 2010. Predicting product adoption in large-scale social networks, in: Proceedings of the 19th ACM international conference on Information and knowledge management, ACM. pp. 1039–1048.

[3] Burges, C.J., 1998. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery 2, 121–167.

[4] Cuadra, L., 2012. Metodología de búsqueda de sub-comunidades mediante análisis de redes sociales y minería de datos. Universidad de Chile.

[5] Domingos, P., Richardson, M., 2001. Mining the network value of customers, in: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, ACM. pp. 57–66.

[6] Eagle, N., Pentland, A.S., Lazer, D., 2008. Mobile phone data for inferring social network structure, in: Social computing, behavioral modeling, and prediction. Springer, pp. 79–88.

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

[8] Kotsiantis, S.B., 2007. Supervised machine learning: a review of classification techniques. Informatica 31.

[9] Ling, R., Yen, D.C., 2001. Customer relationship management: An analysis framework and implementation strategies. Journal of Computer Information Systems 41, 82–97.

[10] Neocleous, C., Schizas, C., 2002. Artificial neural network learning: A comparative review, in: Methods and Applications of Artificial Intelligence. Springer, pp. 300–313.

[11] Ngai, E.W., Xiu, L., Chau, D.C., 2009. Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications 36, 2592–2602.

[12] Palla, G., Derényi, I., Farkas, I., Vicsek, T., 2005. Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814–818.

[13] Payne, A., Frow, P., 2005. A strategic framework for customer relationship management. Journal of marketing 69, 167–176.

[14] Reinartz, W., Krafft, M., Hoyer, W.D., 2004. The customer relationship management process: Its measurement and impact on performance. Journal of marketing research 41, 293–305.

[15] Rokach, L., 2008. Data mining with decision trees: theory and applications. volume 69. World scientific.

[16] Videla-Cavieres, I.F., Ríos, S.A., 2014. Extending market basket analysis with graph mining techniques: A real case. Expert Systems with Applications 41, 1928–1936.