Churn Analysis of a Product of Application Search in Mobile Platform

Gábor SZŰCS, Attila KISS
Inter-University Centre for Telecommunications and Informatics, H-4028 Kassai út 26, Debrecen; Department of Telecommunications and Media Informatics, BME, Hungary
szucs@tmit.bme.hu

Churn is a phenomenon when customers of a subscription based business with regular periodical fees leave the service provider. In this paper churn is investigated from an economic point of view and analyzed for helping the service provider. The investigated provider is a software provider, whose product can be found in a mobile platform, and it is able to search appropriate mobile applications for users. There are different search services, and this application search product is a suitable example to investigate the competitive advantage. The usual usage, the functionalities of the product and the customer's behavior are analyzed. One of the largest contributions of this paper is the extended definition of the churn for services in application market, and the appropriate user identification for this definition, furthermore the suggested solution to predict the churn by data mining methods. More predictive data mining models have been elaborated to forecast the churn of customers, and based on a combined model of them a business solution has been provided with more accurate churn prediction.
Keywords: *Application Search, Churn, Customer Lifetime Value, Data Mining*

1 Introduction

Customer churn is often a rare event in service industries, but of great interest and great value. 0. Generally there are two basic types of churn 0:

- the first is *incidental churn*: this happens when there are sudden changes in the customer's circumstances, and although otherwise he would continue using the service, with the new situation he cannot. Most frequent examples of this type are financial trouble and geographical relocation.
- the second type of churn which is interesting from a business point of view is *deliberate churn*. This happens when customers are dissatisfied with some elements of the service. Most common reasons are the following: poor quality of service, unsuitable experiences with the provider, or higher standards or cheaper alternatives at the competitors. Two types of deliberate churn can be distinguished 0:
    - o *Comparison churn*: when the customer chooses another service provider because his offer or service is better in some aspects.
    - o *Frustration churn*: when the service level falls below the customer's own expectations (also called as the switching barrier) and he decides to leave the provider because of the insufficient service level.

There are many works about customer churn prediction 000, e.g. in broadband internet services 0, in cellular wireless network services 0, in mobile telecommunication industry 0000, in banks 0.

There are three important elements to consider and understand in a customer's interaction with a subscription business 0:

- customer experience: What measurable service experience is the customer exposed to?
- customer perception: How does the customer claim to value the experience?
- customer behaviour: How does the customer actually behave after the experience?

Traditionally the churn is a phenomenon when customers of a subscription based business with regular periodical fees leave the service provider. The investigated provider: Distinction, whose product, AppFlow can be found in mobile platform, and it is able to search appropriate mobile applications for users. This product is free, so we extend the definition of churn, where we consider a user as a churner, when the usage will strongly decrease. This is not so much definite as in traditionally churn, so the identification of users (as possible churners) should be solved. In the next session the churn is investigated from an economic point of view and analyzed for helping the service provider.

2 Churn from Economic Point of View

2.1 Costs of Churn

There are two major costs associated to churn:

- the first is the high *costs of securing new customers* (which gets more expensive with the market maturing and getting more saturated) as opposed to keeping the customers already subscribed.
- the second is the *retention cost*, which is still a lot less, than the costs for getting new customers to sign, but it is a significant cost, especially if the retention strategy is faulty.

At customer retention strategy the value of the customers should be considered. If customer lifetime value or *customer equity* is high, then probably the final *customer profitability* after the successful retention campaign will be higher as well, thus the company's profit level will be higher 0.

2.2 Customer Lifetime Value

Customer equity is all the income that the customer has produced for the service provider since he first signed. Profitability is all the income minus all the costs that came with getting the customer sign in the first place and stay since up until the present time. *Customer lifetime value* (LTV) is another profit centered customer value measure, as can be seen in equation (1), where R_i is the income that customer generated, C_i is the costs that came with the customer during an *i* period of time, *d* is the expected rate of return, so the denominator is the component responsible to calculate the *net present value* for the whole sum.

$$LTV = \sum_{i=0}^{n} \frac{(R_i - C_i)}{(1+d)^{i-0.5}} \tag{1}$$

Taking the customer loyalty into account the *customer lifetime value* could be calculated not only for the past, but for the future as well. This would give more realistic formula, as can be seen in equations (2)-(4) based on the work of 0, where equation (2) concerns to past, equation (3) concerns to future, and equation (4) is the total *customer lifetime value* in time *i*.

$$_{P}LTV_i = \sum_{t_i=0}^{N_i} G_P(t_i)(1+d)^{N_i - t_i} \tag{2}$$

The $_{P}LTV_i$ is the sum that provider gets all the profits until N_i (as present time) as can be seen in equation (2), where $G_P(t_i)$ is the profit in time period t_i (in the past), and *d* is the expected rate of return.

If we could know the probability of churn of a given customer, then we could estimate the length of stay of the customer before churning. Using this information we are able to calculate the sum of (supposed) expected future profits as can be seen in equation (3), where t_i is time index, N_i is time spent with the provider until now, E(i) is length of expected stay of the customer before churning, $G_F(t_i)$ is expected future profit contribution, and $B(t_i)$ is potential customer (non financial) benefits for the time period 0.

$$_{F}LTV_i = \sum_{t_i=N_i+1}^{N_i+E(i)+1} \frac{G_F(t_i) + B(t_i)}{(1+d)^{t_i - N_i}} \tag{3}$$

$$LTV_i = {}_{P}LTV_i + {}_{F}LTV_i \tag{4}$$

2.3 Churn Management Strategy

To build a successful customer churn prediction model, a classification algorithm should be chosen that fulfills two requirements: strong classification performance and a high level of model interpretability 0 0.

There are three major information extraction tasks for a company to be able to use data to build an efficient churn management strategy:

- understanding customer experience: what measurable service experience is the customer exposed to?
- familiarizing ourselves with customer perception: how does the customer evaluate our services?
- monitoring customer behaviour: how does the customer behave after using the service?

Customer churn prediction is based on the assumption that an unknown dependence relationship exists between churn variable y and customer information variable x. The usual method for building a churn prediction system, executing the prediction itself and retention strategies is as follows 0:

1. building IT infrastructure, procuring the necessary data;
2. defining scope for the analysis,
3. exploratory data analysis and identification of best data,
4. data preprocessing and semantics,
5. variable/feature analysis and selection,
6. validation of the selected data fields,
7. making the model and running it,
8. testing the results and making a new model if necessary,
9. evaluating of the final results.

3 Identification of Users with Segmentation

3.1 Problem and the Solution

In traditional services the customers of a subscription based business with regular periodical fees are loyal users until they leave the service provider. But in a search product the loyal users are not defined as clearly, identification of them is a hard problem. However in order to analyze the churn, we should distinguish the users that have only tried the search product (1^{st} group) from the users that have frequent used this (2^{nd} group). The churn phenomenon can occur among the 2^{nd} group of users, when persons 'leave' the search product and do not use it anymore.

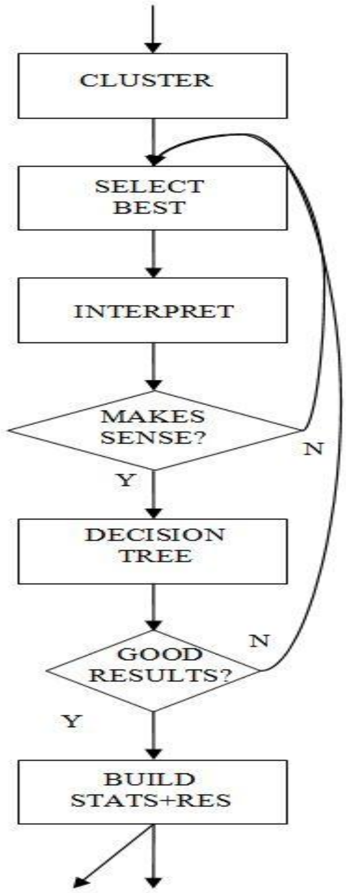

Fig. 1. Multistep segmentation for identification of users

We have solved the user identification problem by multistep segmentation as can be seen in Figure 1. The multistep segmentation contains data mining part and business part as well. The first step is the clustering (CLUSTER in Figure 1), where more clustering algorithms can be tried. The second step is the selection (SELECT BEST in Figure 1), where the best clustering algorithm is selected based on mathematical evaluation. The next step is the interpretation (INTERPRET in Figure 1), where a human tries to interpret the clusters from business point of view. If this interpretation is successful, i.e. clusters have business sense (MAKES SENSE in Figure 1), then the flow of the multistep segmentation can be continued, otherwise the procedure goes back to the second step. The final idea for evaluation is a classification step; decision tree (DECISION TREE in Figure 1) is able to measure the distinguish ability of the users, which helps with measurement the goodness of segmentation.

Based on the result of the classification step the flow is continued (GOOD RESULTS in Figure 1) or goes back to the second step. At the final step the segmentation of the users are available, statistics can be built (BUILT STATS + RES in Figure 1) and based on final result the users can be selected for further (churn) analysis. These steps of multistep segmentation for user identification problem are discussed below.

3.2. Clustering

For clustering the variables a distance function has to be defined. Some variables from most important statistics are listed below:

- *start*: is the time and date of starting the session.
- *liveid*: is the identifier of the user.
- *end*: is the time and date of finishing the session.
- *freeAppsDownloaded*: is the number of free applications downloaded during the session.
- *paidAppsDownloaded*: is the number of paid applications downloaded during the session.
- *eventType_X*: is the number of occurrences of an event (possessing event type X) during the session.
- *p_swipesUp_X*, *p_swipesDown_X*: Number of swipes up and down in a session corresponding to the user.
- *p_clickCount_X, p_numberOfResults_X, p_SearchQuery_X*: indicators about clicking and results of search

Furthermore there are lots of variables about the clicking, user actions, searching and results of searching. All variables (169) have been used in clustering. In order to get a large set results (in this case partitions) more clustering algorithms have been used, like k-means with different parameters, and other clustering methods with two steps and different parameters.

3.3 Selection the Best Result and Interpretation

The mathematical evaluation consists of statistical viewpoints, indicators, like the number (and the ratio) of the elements in different clusters, the maximal and average value of distances between each element and the central point.

After the filtering the possible adequate results, in the interpretation step we have interpreted the clusters from business point of view:

- cluster 1: users, who have tried the search product only once; they (*trialing users*) have not spent too much time with the usage of the product and they have not downloaded mobile applications. They are 36% ratio of all users and their common parameters are:
  - average 1 minute usage
  - zero downloading
- cluster 2: users, who use the search product many times, but they (*browsing users*) download only very few mobile applications (typical number of downloads is one or zero per login). They are 24% ratio of all users and their common parameters are:
  - average 4.5 minutes usage
  - average 0.4 application downloads / login
- cluster 3: users, who use the search product many times, and they (*active, paying users*) are willing to pay for high quality mobile applications. They are 16% ratio of all users and their common parameters are:
  - average 4.5 minutes usage
  - 0-2 free application downloads / login
  - average 0.5 paid application downloads / login
- cluster 4: users, who use the search product numbers of times and like to try everything, but they (*active, free users*) are not willing to pay for high quality mobile applications. They are 24% ratio of all users and their common parameters are:
  - average 5 minutes usage
  - average 1.6 free application downloads / login

This interpretation was successful, i.e. clusters have business sense. In a classification task these clusters have been considered as classes, and decision tree has been built (at

70% training and 30% validation dataset separation) for evaluation of goodness of segmentation. The accuracy of classification has been more than 0.9, which shows an adequate result of the multistep segmentation procedure.

4 Churn Prediction with Data Mining
4.1 Churn Prediction

Based on the user segmentation described in the previous session the cluster 3 and cluster 4 can be considered active users, so in churn analysis these type of users are the most important ones.

Now AppFlow's search service is free of charge at the moment, so the churn cannot be interpreted in the traditional way, but a similar idea can be applied. Churn is basically the customer deciding not to use the service anymore. Since AppFlow is free, the following criteria are constructed for own churn definition:

- if the user does not count as a subscriber (in cluster 1), then churning does not make any sense.
- if the browsing and active user (in cluster 2, 3, and 4) has logged in more than 90 days ago, he/she is considered a churner. (If somebody does not use a service for 30 days, that is considered normal in casual usage, 60 days is a questionable point, but at 90 days it is quite certain that the user will not use the program any more.)

Based on half year statistical data the ratio of users in cluster 1 and all users is 13.3%, furthermore 9.7% was found to be churners based on own definition in the original dataset and the rest (77%) are non churner users. The cluster 1 is excluded from further examinations (the users in cluster 1 are deleted from the database). This means that the dataset contains 107465 customers of which 11.2% are churners and 88.8% are non churners.

The churn has been analyzed on approximately 500 million log entries about user sessions in a csv file originally come from Distinction's database. The following preprocessing and data manipulation preparations have been used:

- converting the file to a format that SPSS could properly read as input,
- filling up incomplete data or correcting misplaced and wrong values,
- deleting outlier and faulty or unrepairable data,
- dealing with the date manipulation options that SPSS provides,
- creating new date and usage related variables.

For churn prediction different data mining algorithms can be used. To find the best one in our explicit case, it is best practice to try a few models and then compare the success measures of the given results. We can use the ROC curve and the AUC (area under curve) as measurements, or the Gini coefficient which can be calculated based on the AUC, but in this paper we have used accuracy, i.e. (TP+TN)/all, where TP means true positive and TN means true negative. In the next session the data mining algorithms are described that have been used for churn prediction 0.

4.2 Usable Data Mining Algorithms

Naïve Bayes classification calculates the probability that a given input sample belongs to a certain class. Given a sample X which consist of a feature vector $\{x_1, \dots, x_n\}$ the probability for class y_j can be obtained by the following:

$$P(y_j \mid X) = P(X \mid y_j) \cdot P(y_j) / P(X) \quad (5)$$

where $P(y_j)$ is the prior probability of y_j and P(X) is constant. However Naïve Bayes assumes that the conditional probabilities of the independent variables are statistically independent, the likelihood can be described by:

$$P(X \mid y_j) = \prod_{i=1}^{n} P(x_i \mid y_j) \quad (6)$$

Considering that there are a number of possible classes the unknown sample X is classified based on the following:

$$c = \arg\max_{y_j \in Y} P(y_j \mid X) \qquad (7)$$

We can use that the P(X) is constant and the conditional probability is equal to product of the independent conditional likelihoods 0:

$$c = \arg\max_{y_j \in Y}\left(P(y_j) \cdot \prod_{i=1}^{n} P(x_i \mid y_j) \right) \qquad (8)$$

Random Forests is an algorithm, which uses multiple trees as a voting system and the final decision is based on majority vote. Classification with this method is useful, because it is accurate and very efficient on large databases, it can handle many input variables and cases and it is very transparent. This transparency comes from the set rules of the trees that we can afterwards use and also means that on the same dataset the model is valid for reuse without calibration 0.

SVM (*Support Vector Machines*) operates by constructing an N-dimensional hyper plane, which optimally separates the data into two categories – in case of a discrete two dimension problem like churn. A SVM classifier can be trained by finding a maximal margin hyper plane in terms of a linear combination of subsets of the training set. If the input feature vectors are nonlinearly separable, SVM firstly maps the data into a high dimensional feature space by using the kernel trick, and then classifies the data by the maximal margin hyper-plane 000.

5 Applied Method and its Results in Churn Prediction

5.1 Input Data Models

We have constructed three input models for every data mining method. The idea behind this was that we will also be able to examine which information group is necessary or enough for the prediction. The three input models have used the following variables:

- first input data model: the 4 spent_time variables and the two apps_downloaded variables: *freeAppsDownloaded* and *paidAppsDownloaded*.
- second input data model: the factored variables as results of factor analysis
- third input data model: in addition to the first input model the raw usage variables from the original aggregation are added

At the first input data model only some aggregate variables are defined: a sum variable for the time spent and the number of sign ins, and additional two variables about time spent. The importance of these from a business point of view is as follows: The mean of the variables shows the usual service usage for the subscribers, about what features they prefer, what parts of the service are they using mostly. The time spent, the apps downloaded and the number of sign ins are the most important variables from a business point of view. They show the main characteristics of the usage: whether the client has spent a lot of time in the program, and whether he found useful results or not.

At the second input data model a factor analysis is used to decrease the number of variables. The PCA has been used at for 212 dimension data, and at the end of the dimension reduction algorithm 39 factor variables remain.

At the third input data model in addition to the first input model the raw usage variables from the original aggregation are added, e.g. number of swipes up and down in a session corresponding to the user, indicators about clicking, and the number of occurrences of other events during the session.

5.2 Applied Data Mining Algorithms for Churn Prediction

SPSS is a well known and widely used industrial solution in data mining, and the next algorithms have been used in SPSS for own churn prediction task.

Logit regression is a type of analysis used for outcome prediction of a categorical dependent variable based on a number of predictor variables. The probabilities describing the possible outcome of a single trial are modeled, as a function of explanatory variables, using a logistic function. Logistic regression measures the relationship between a categorical dependent variable and usually continu-

ous independent variables, by converting the dependent variable to probability scores.
By this algorithm the overall successful prediction rates can be seen in Figure 2 with different input data models described above. The best accuracy has been 89.4% at the input data model 2 and 3.

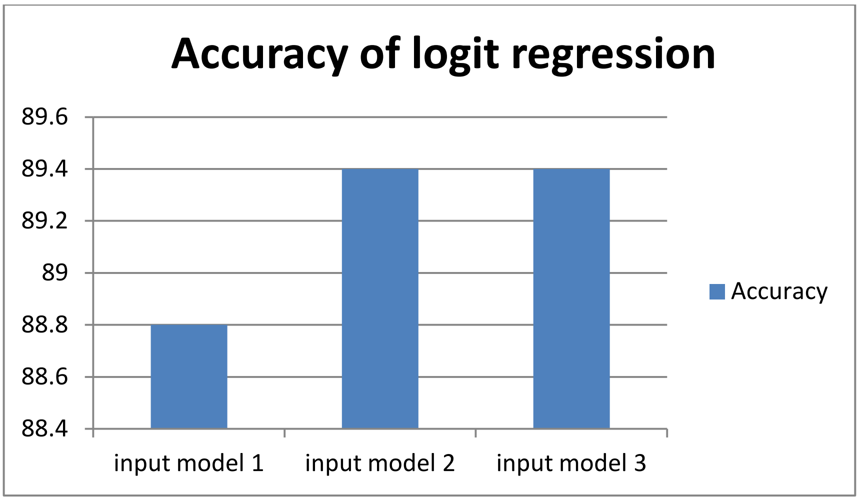

Fig. 2. Accuracy of logit regression

Multilayer Perceptron Network (MLP Network) is a feed-forward artificial neural network. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node possesses a nonlinear activation function and an MLP can distinguish data that are not linearly separable. We have used MLP only two times: for input data model 1 and 2, because at input data model 3 the execution was too slow to wait the results because of huge amount of independent variables. After running the algorithm, the overall success measure was 76.8% at the input data model 1 and 77.2% at data model 2 as can be seen in Figure 3.

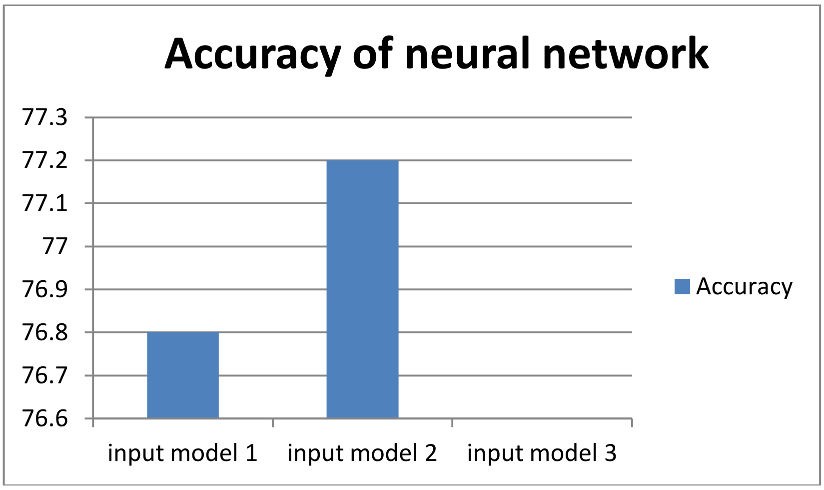

Fig. 3. Accuracy of neural network with multilayer perceptron

Radial basis function (RBF) network is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. This is similar to MLP, because of networks structure. The RBF network has been faster in our data set, so three runs can be executed. The larger accuracy has been 88.3% at input model 1 as can be seen in Figure 4.

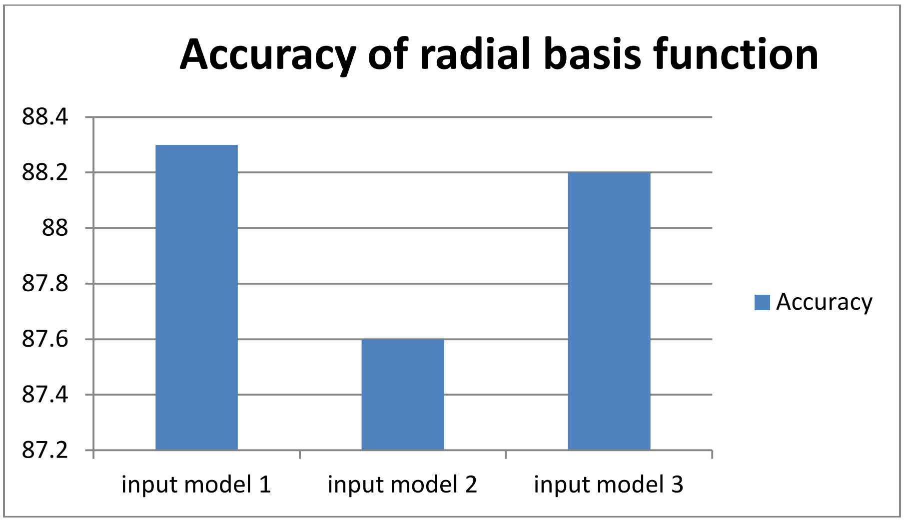

Fig. 4. Accuracy of neural network with radial basis function

Decision trees work on the basic idea of divide and conquer, building a tree. Initially the method starts to search for an attribute with the best dividing variable (e.g. information gain) and divides the tree into sub-trees at the root node. The algorithm produces this recursively until the termination conditions are met. Once the tree is created, rules can be obtained by traversing each branch of the tree. This algorithm usually performs very well with churn related problems, and it is also a favorite because the rules constructed are very easy to understand furthermore this can be used for further analysis. The most commonly used types of decision tree algorithms are as follows: ID3, C4.5, C5.0 (e.g. in 0), CART, CHAID.

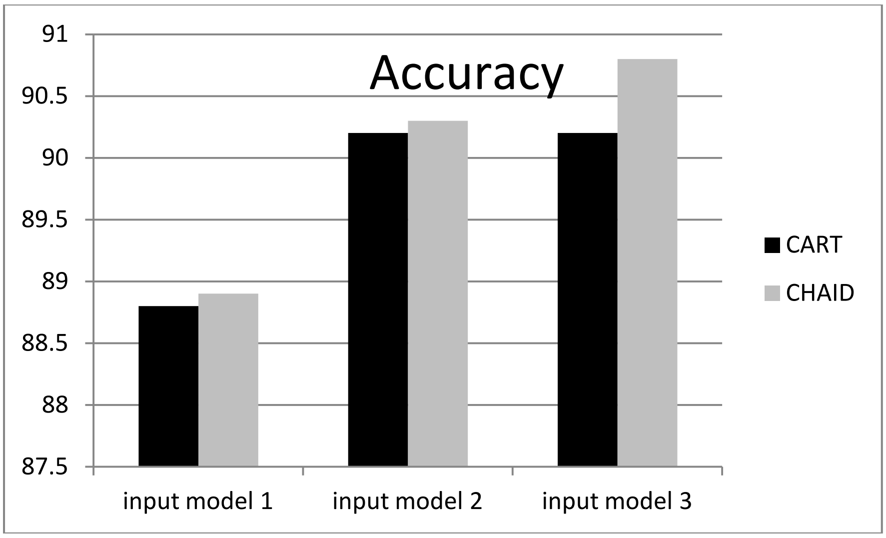

Fig. 5. Accuracy of different decision trees

The best accuracy in own churn prediction task has been at the last input data model: 90.8% with CHAID as can be seen in Figure 5.

KNN (*K Nearest Neighbors*) is a classification algorithm that works based on distances among training set data. Once it has the training data, it goes through the data waiting to be classified and it determines the class for each case based on the k nearest training cases class. The determination is a simple majority vote. By this algorithm the overall successful prediction rate, i.e. the accuracy has been 88.8%.

5.3 Combination of the Algorithms

It is a best practice to not just use the best classifier, but choose the best three or best five, run classification and predictions with all 3 or 5 of them, and then use the results for a majority vote, thus determining the churn prediction with more than one algorithm. The advantage of this method is that this way we can combine the strength of the selected methods and we can protect against blind spots of algorithms.

Using the best three algorithms described above, a simple majority voting algorithm has been implemented. The combined success measure (accuracy) has been 91.3%, which means that with the methods combined we actually managed to gain a 0.5% increase compared to the highest individual success factor.

6 Conclusion

Data mining methods are really useful and well performing in churn analysis and management. One of the largest contribution of this paper is the extended definition of the churn for services in application market, and the appropriate user segmentation for this definition, furthermore the suggested solution to predict the churn by data mining methods. It is worth to try more than one approach empirically and to choose the best one or best ones. Combining methods can be very beneficial, because the methods can take out each other's weaknesses if properly planned. It is important to check the results and always keep in mind the business goal and the business interpretation.

There is a possibility to further experiment with different models at hand. The possible settings for the algorithms are endless, so fine tuning the results is always a possibility. There are many data mining algorithms, but there are few of them, which help to understand the business aspects as well. For example the decision tree is adequate to interpret the data from a business point of view. This would explain why users are churning, what features or properties of the product need to be changed.

Acknowledgements

The publication was supported by the TÁMOP 4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

References

[1] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, Computer Assisted Customer Churn Management: State-Of-The-Art and Future Trends. Computers & Operations Research. 34(10) October, 2007, pp. 2902-2917.

[2] M. K. Kima, M. Ch. Park, and D. H. Jeonga, The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. Telecommunications Policy. 28, pp. 145–159, 2004.

[3] D. A. Kumar and V. Ravi, Predicting credit card customer churn in banks using data mining. Int. J. Data Analysis Techniques and Strategies. Vol. 1, No. 1. 2008.

[4] W. Verbeke, K. Dejaeger, D. Martens, J. Hur and B. Baesens, New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. European Journal of Operational Research. 218, pp. 211–229, 2012.

[5] K. Morik and H. Köpcke, Analysing Customer Churn in Insurance Data – A Case Study. PKDD. LNAI 3202, pp. 325–336, 2004.

[6] H. Hwang, T. Jung and E. Suh, An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. Expert Systems with Applications. 26, pp. 181–188, 2004.

[7] B. Huang, M. T. Kechadi and B. Buckley, Customer churn prediction in telecommunications. Expert Systems with Applications. 39, pp. 1414–1425, 2012.

[8] Y. Xie, X. Li, E. W. T. Ngai and W. Ying, Customer churn prediction using improved balanced random forests. Expert Systems with Applications. 36, pp. 5445–5449, 2009.

[9] S. Kim, K. Shin and K. Park, An Application of Support Vector Machines for Customer Churn Analysis: Credit Card Case. ICNC 2005, LNCS 3611, pp. 636–647.

[10] M. Kon, Customer Churn. Stop before it starts, Mercer Management Journal (MMJ), 17, pp. 54-60, 2004.

[11] M. K. Kim, M. C. Park and D. H. Jeong, The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. Telecommunications policy, 28(2), 145-159, 2004.

[12] A. D. Kumar and V. Ravi, Predicting credit card customer churn in banks using data mining. International Journal of Data Analysis Techniques and Strategies, 1(1), pp. 4-28, 2008.

[13] G. E. Xia and W. D. Jin, Model of customer churn prediction on support vector machine. Systems Engineering-Theory & Practice, 28(1), pp. 71-77, 2008.

[14] J. H. Ahn, S. P. Han and Y. S. Lee, Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. Telecommunications Policy, 30(10), pp. 552-568, 2006.

[15] S. Y. Hung, D. C. Yen and H. Y. Wang, Applying data mining to telecom churn management. Expert Systems with Applications, 31(3), pp. 515-524, 2006.

[16] P. C. Pendharkar, Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services. Expert Systems with Applications, 36(3), 6714-6720, 2009.

[17] Y. F. Wang, D. A. Chiang, M. H. Hsu, C. J. Lin and I. Lin, A recommender system to avoid customer churn: A case study. Expert Systems with Applications, 36(4), 8071-8075, 2009.

[18] C. F. Tsai and Y. H. Lu, Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10), 12547-12553, 2009.

[19] Z. Jamal and R. E. Bucklin, Improving the diagnosis and prediction of customer churn: A heterogeneous hazard modeling approach. Journal of Interactive Marketing, 20(3), 16-29, 2006.

[20] R. K. Gopal and S. K. Meher, Customer churn time prediction in mobile telecommunication industry using ordinal regression. In Advances in Knowledge Discovery and Data Mining, Springer Berlin Heidelberg, pp. 884-889, 2008.

[21] Y. Zhao, B. Li, X. Li, W. Liu and S. Ren, Customer churn prediction using improved one-class support vector machine. In Advanced Data Mining and Applications, Springer Berlin Heidelberg. pp. 300-306, 2005.

[22] B. Q. Huang, M. T. Kechadi and B. Buckley, Customer churn prediction for broadband internet services. In Data Warehousing and Knowledge Discovery, Springer Berlin Heidelberg. pp. 229-243, 2009.

[23] I. Bose and X. Chen, Hybrid models using unsupervised clustering for prediction of customer churn. Journal of Organizational Computing and Electronic Commerce, 19(2), pp. 133-151, 2009.

[24] J. Burez and D. Van den Poel, Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3), pp. 4626-4636, 2009.

[25] K. W. De Bock and D. Van den Poel, Reconciling performance and interpretability in customer churn prediction using ensemble learning based on generalized additive models. Expert Systems with Applications, 39(8), pp. 6816-6826, 2012.

[26] W. Verbeke, D. Martens, C. Mues and B. Baesens, Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Systems with Applications, 38(3), pp. 2354-2364, 2011.

Gábor SZŰCS was born in 1970 in Hungary. He has received MSc in Electrical Engineering from Budapest University of Technology and Economics (BME) in 1994. He is experienced in modeling and simulation, railway systems, traffic systems; he has received PhD degree in this field from BME in 2002. His further and currently research areas are data mining in business areas, multimedia mining, content based image retrieval, semantic search. He is associate professor at Department of Telecommunications and Media Informatics of BME. The number of his publications is more than 80. He is vice president of the Hungarian Simulation Society (EUROSIM), deputy director of the McLeod Institute of Simulation Sciences Hungarian Center. He has earned János Bolyai Research Scholarship of the Hungarian Academy of Science in 2008.

Attila KISS was born in Hungary, in 1990. He received his B.Sc. degree in computer engineering from the Budapest University of Technology and Economics, Hungary in 2012, with a substitute degree in IT entrepreneurship from Aquincum Institution of Technology, Budapest, Hungary in same year. He is currently concluding M.Sc. studies in business information systems at the Budapest University of Technology and Economics, Hungary, with a term abroad at University of Technology Sydney, Australia in 2012. In 2011 he was a trainee at SAP Labs Hungary, working on a new wave SAAS development project. He was a board member at Simonyi College for Advanced Studies with the role of CFO in 2010, and President in 2011. His main research interests are revolving around customer segmentation and churn management. Mr. Kiss awards and honors include the Life Technologies Scholarship (Life Technologies Inc, 2011) and the UTS Business Faculty Scholarship (University of Technology Sydney, 2012).