DE GRUYTER OPEN

*Folia Oeconomica Stetinensia*
DOI: 10.1515/foli-2015-0006

WYDZIAŁ NAUK EKONOMICZNYCH I ZARZĄDZANIA
UNIWERSYTET SZCZECIŃSKI

# HYBRID C&RT-LOGIT MODELS IN CHURN ANALYSIS

Mariusz Łapczyński, Ph.D.

*Cracow University of Economics*
*Department of Market Analysis and Marketing Research*
*Rakowicka 27, 31-510 Cracow, Poland*
*lapczynm@uek.krakow.pl*

**Received 14 April 2014, Accepted 24 October 2014**

## Abstract

This article attempts to explain and predict the termination of relationships in telecommunications services by using the hybrid C&RT-logit model. The combination of decision trees (C&RT algorithm) with the logistic model enriches the model interpretation and sometimes improves the accuracy of prediction. Decision trees permit to detect interactions among variables and make the model resistant to outliers and to lack of data. On the other hand, the logistic model can extend the interpretation by using odds ratios. The solution delivered by the hybrid approach was compared with the decision tree model and the logistic model. Due to the difficulty in obtaining the real dataset from the Polish market, it was decided to build a model based on the data obtained from the repository http://www.dataminingconsultant.com/DMMM.htm. The models' performance was estimated by using popular measures such as accuracy, recall, precision, true negative rate, G-mean, F measure and lift charts.

**Keywords:** churn analysis, hybrid C&RT-logit model.

**JEL classification:** C35, M31.

## Introduction – churn analysis and relationship marketing[1]

Churn analysis appears in relationship marketing literature in the context of customer retention (model ACURA – acquisition, cross-sell, up-sell, retention, advocacy and customer life cycle), satisfaction and loyalty of customers. Satisfaction treated as an unobservable construct determines the retention, which is an observable variable[2]. It seems that the lack of consumer satisfaction in a natural way explains their defection, but it is not always the case. There are several possible reactions of dissatisfied consumers[3]:

- exit – customers cease to purchase the company's products, they can also start to cooperate with competitors,
- voice – customers report critical findings concerning the offer and demand to solve the problem,
- loyalty – customers remain loyal to the supplier, which is caused by its perceived or actual monopolistic position, continued cooperation may also be caused by high switching costs, ideological reasons or inertia,
- collaboration – customers try to solve the problem by collaboration with the company, it helps to strengthen the ties between them and increases the value (benefits) for both parties.

The connection between satisfaction and loyalty, and the need of retention is not always obvious. One draws attention[4] to the fact that a satisfied customer can be unprofitable for the company due to the high service costs. A satisfied customer can also be a "happy slave" who is not aware of the existence of competitors' offer. It may be impossible to retain him when he realizes it.

Customers' dissatisfaction is only one of several possible reasons for the termination of relationships. Defection can be also caused by[5]:

- satiation – consumers are bored with the current product or service and look for new solutions,
- competitor's offer – customers perceive an alternative offer as more attractive one;
- conflict with the supplier,
- high exit barriers – if the buyer does not have the possibility of exit (or it is very difficult to do it), this can increase his discomfort and increase the probability of defection at the earliest opportunity.

It is widely recognized that the costs of customer acquisition are higher than the costs of retention, which in the case of certain services becomes measurable. On average, the profitability of bank clients is significantly higher after 5 years from opening the account. In the case of automobile insurance policies clients become profitable after 7 years[6]. ROI (return on investment) is sometimes up to 10 times higher in retention strategies than in acquisition strategies. The cost of acquiring new customers can be from 5 to 12 times higher than the cost of maintaining the existing ones[7].

An increased emphasis on customer retention can cause[8]:

- reduction in costs of acquiring new customers (a high retention rate makes the potential number of new customers lower),
- increase in sales volume and value for existing customers,
- reduction in service costs of existing clients,
- word-of mouth recommendations by the retained customers,
- lower sensitivity of the existing clients to premium pricing,
- problems for new companies to enter the market or increase the market share.

Churn analysis means a partial or absolute termination of relationships. To describe that group of predictive models one uses the above-mentioned term “churn analysis” (usually in the telecommunications industry), the term “attrition analysis” (usually in the financial services market) or the term “retention/defection analysis” (in the marketing and CRM literature). Churn rate is considered a key measure of customer loyalty. This means the percentage of customers who decided not to purchase company’s products or services. It is usually calculated per one year, quarter or month. If there is contractual agreement between parties (e.g. mobile phone subscription, magazine subscription, contracts with internet service providers), churning is not renewing the contract for a further period. In non-contractual settings churning can be estimated on the basis of a sales analysis or questionnaire research[9]. Apart from the churn rate one can estimate the retention rate, which represents the percentage of customers prolonging cooperation in the group of clients for whom the contract expires[10].

Churn analysis was primarily used in the telecommunications services[11] and in the financial services market[12]. One can observe, however, an increasing number of churn models in different areas, such as retail trade[13], insurance market[14], paid-TV subscription[15], magazine subscription[16] and educational services[17]. This article attempts to explain and predict the termination of relationships in telecommunications services by using the hybrid C&RT-logit model.

## 2. Hybrid predictive C&RT-logit model

While building churn models one commonly uses decision trees, logistic regression, random forest, boosted trees or support vector machines. Recently, however, some researchers have tried to combine different analytical tools, which is called a hybrid approach. One can easily find a combination of the cluster analysis with decision trees, genetic algorithms with neural networks or decision trees with logistic regression.

The hybrid C&RT-logit model used in this study combines C&RT (classification and regression trees) algorithm with logistic regression. The STATISTICA software was used for the data analysis and therefore the abbreviation CART (classification and regression trees) that is a registered trademark of Salford Systems company was replaced with the acronym C&RT.

CART is considered as one of the most advanced decision tree algorithms[18]. The dependent variable and the independent variables can be measured on any type of scale (nominal, ordinal, interval, ratio), and practically no assumptions have to be met. The only technical limitation may concern the number of predictors or the number of categories of qualitative variables. Logistic regression models are very popular in the case of a binary dependent variable and a set of independent variables at any level of measurement.

The features of the CART algorithm that distinguish it from the logistic regression model are as follows:

- automatic selection of the best predictors (it also builds the importance ranking of independent variables),
- no need for the transformation of variables (e.g. logarithm, square root),
- automatic detection of interaction effects,
- resistance to outliers,
- utilizing surrogate variables while classifying cases with missing data,
- minimal supervision of the researcher is required while building the model.

CART algorithm detects the data structure, but in the case of trees with many leaves it does not provide a clear presentation of the model. It can also happen that a large number of terminal nodes represents very simple relationships between the variables.

The construction of logistic regression models, in turn, requires the supervision of an experienced analyst and frequently takes much longer than the construction of the decision tree. Logit models are sensitive to outliers and require imputation of missing data (cases with missing data are removed from the analysis). The big advantage of this approach is the ability to calculate the unique probability of belonging to a class (category dependent variable) for each case. On the other hand, decision trees provide as many probabilities as many leaves they have for each terminal node and cases belonging to it.

The combination of decision trees (CHAID algorithm) with logistic regression carried out by Lindahl and Winship was probably the first attempt to build this kind of a hybrid model[19]. Hybridization was based on the sequential use of these analytical tools. After the initial exploration of dataset by using CHAID algorithm cases were divided into terminal nodes. In the second step of the procedure a separate regression model was built for each leaf.

Another concept of hybridization was proposed a few years later[20]. It combined decision trees (CART algorithm) with logit models. This time it was also a two-step procedure, however, the set of independent variables in the logit model was supplemented with an additional variable whose categories informed about the terminal node to which the case was assigned. The new variable was transformed into a set of dummy variables. CART model from the first stage of the procedure was based on the same set of independent variables, and each leaf took into account the interaction between the predictors. The authors pointed out that such hybridization is more effective, because the partition of the dataset into subsets according to the first concept is connected with a reduction of sample size (instances are divided into terminal nodes) and the loss of information (it can happen that the sets of independent variables for each logit model will be slightly different). Moreover, patterns discovered by logit models are local (limited to the leafs) and the variability of the dependent variable and variance of predictors is lower in terminal nodes than in the entire dataset.

The advantages of the CART-logit approach include a higher predictive accuracy of the hybrid model, a faster detection of interactions by the CART algorithm and, in general, no need to replace missing data. Authors distinguished several ways of dealing with missing data while building the CART-logit model. One can ignore the problem because CART algorithm can perform well anyway. Analysts can assign CART-predicted probabilities to these cases or assign hybrid-predicted probabilities to remaining cases. One can also impute missing data or add a dummy variable indicating which case has missing value.

## 3. An attempt to build hybrid CART-logit model in churn analysis

### 3.1. Description of dataset

The dataset used in this experiment refers to churn analysis. Initially, the intention of the author was to build a model relating to the Polish market. Unfortunately, the problem with obtaining and publicizing such data resulted in using data from popular online repositories. The dataset used in this experiment is, according to the author's knowledge, one of the best described and available without any restrictions. It was obtained from www.dataminingconsultant.com/DMMM.htm. The set of observations contains 5,000 cases, the binary dependent variable (churn/no churn) and 20 independent variables. The percentage of churners is equal to 14.14%, which indicates a class imbalance problem. In the first step the dataset was divided into a learning sample (69.9% of the entire dataset) and a test sample (31.1%). The sample size and distribution of the target variable are shown in Table 1 Due to the problem of imbalanced classes random under-sampling was utilized. It means that the majority class (non-churners) in the learning sample was randomly reduced to the level of 70%. The structure of the test sample remained unchanged. For the purposes of the logistic regression model categorical variables have been replaced by dummies.

Table 1. Sample size and distribution of dependent variable

| Dataset | Sample size | Number and percentage of churners in dependent variable |
|---|---|---|
| Entire dataset | 5,000 | 14.4% |
| Learning sample | 3,494 | 506 (14.48%) |
| Learning sample (random under-sampling) | 1,687 | 506 (29.99%) |
| Test sample | 1,506 | 201 (13.35%) |

Source: own calculations.

### 3.2. C&RT model

While building the decision tree model (C&RT algorithm) equal misclassification costs, estimated a priori probabilities and the minimum size of the leaf at the level of 5% of the learning sample (85 cases) were used. Figure 1 shows the model tree, which indicates that:

- if "total day minutes" >246.6, then the probability of churn is equal to 0.718 (leaf ID 3);
- if "total day minutes" ≤246.6 and the "number customer service calls" >3.5, then the probability of defection equals 0.721 (leaf ID 5);
- if "total day minutes" ≤246.6 and the "number customer service calls" ≤3.5 and the "international plan" = 1 (yes), then the probability of churn is equal to 0.608 (leaf ID 6).

Moreover, the variable importance ranking indicates that predictors with a high discriminating power are the "international plan" (100 points), the "number customer service calls" (77), the "total day minutes" (61) and the "total day charge" (61).

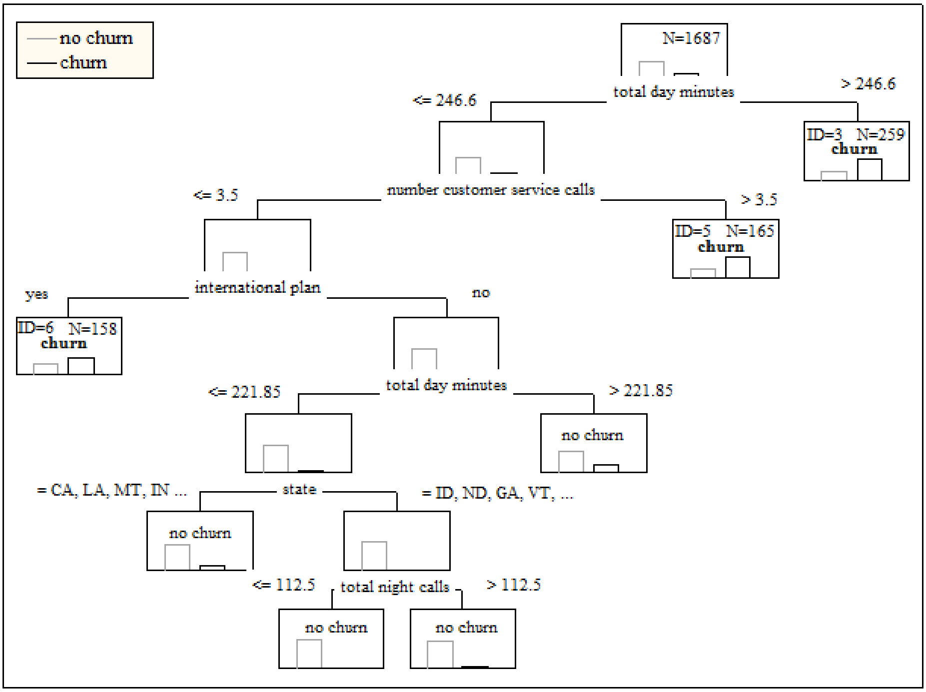


Fig. 1. Decision tree model (C&RT)

Source: own elaboration by using STATISTICA 10.0.

### 3.3. Logistic regression model

The logistic model includes 13 independent variables with significant p-values (8 qualitative marked with the letter "q" and 5 continuous marked with the letter "c") – see details in Table 2. The independent variables that have high positive effects with respect to the response (that increase the probability of churn) are:

- "international plan" (q) – the probability of churning among customers with the international calling plan is almost 12 times higher than the probability among customers without that plan,
- "number for customer service calls" (c) – for every call increase in "number…" the odds of churning increase by approximately 67%,
- "state NJ" (q) – clients living in the state of New Jersey have 2 times greater odds of churning than clients from the states not included in the model,

– “state MT” (q) – customers from Montana have 3.65 times greater probability of churning than customers from the states not included in the model,
– “state CA” (q) – probability of churning among buyers from California is about 6 times higher than the probability among buyers from the states not included in the model.

Table 2. Results of logistic model

| Variable | Estimate | Standard error | p value | Odds ratio |
|---|---|---|---|---|
| Intercept | –5.594 | 0.443 | 0.000 | 0.00 |
| account length (c) | 0.003 | 0.002 | 0.046 | 1.00 |
| international plan (q) | 2.466 | 0.192 | 0.000 | 11.78 |
| voice mail plan (q) | –1.111 | 0.173 | 0.000 | 0.33 |
| total day minutes (c) | 0.014 | 0.001 | 0.000 | 1.01 |
| total eve minutes (c) | 0.005 | 0.001 | 0.000 | 1.00 |
| total intl calls (c) | –0.077 | 0.026 | 0.004 | 0.93 |
| number customer service calls (c) | 0.511 | 0.046 | 0.000 | 1.67 |
| state NJ (q) | 0.773 | 0.358 | 0.031 | 2.17 |
| state RI (q) | –1.738 | 0.721 | 0.016 | 0.18 |
| state MT (q) | 1.295 | 0.417 | 0.002 | 3.65 |
| state VA (q) | –1.364 | 0.613 | 0.026 | 0.26 |
| state IL (q) | –2.911 | 1.101 | 0.008 | 0.05 |
| state CA (q) | 1.806 | 0.589 | 0.002 | 6.09 |

Source: own elaboration by using STATISTICA 10.0.

### 3.4. The hybrid C&RT-logit model

Prior to building the hybrid model, the size of decision tree was reduced to four terminal nodes (Figure 2). In leaf ID 3 there are customers for whom the daily number of minutes of calls is higher than 246.6. The leaf ID 5 includes customers for whom the “total day minutes” is fewer than or equal to 246.6 and the number of calls to the call center exceeds 3 The leaf ID 6 consists of buyers having international plan calls for whom the daily number of minutes is fewer than or equal to 246.6 and the number of calls to the call center is fewer than or equal to 3. In the terminal node ID 7 there are customers that do not have international plan, for whom the daily number of minutes is fewer than or equal to 246.6 and the number of calls to the customer service center is fewer than or equal to 3.

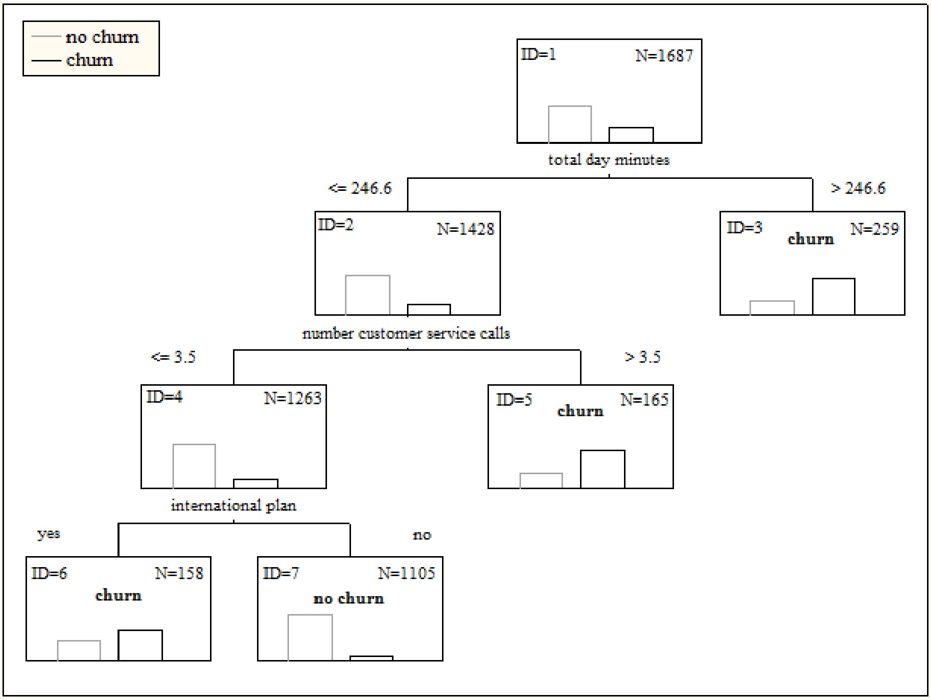


Fig. 2. Reduced decision tree model

Source: own elaboration by using STATISTICA 10.0.

The hybrid model was supplemented with an additional variable "terminal node", which was transformed to dummies with the reference category "leaf ID 3". Table 3 presents the results of the hybrid approach.

Table 3. Results of hybrid C&RT-logit model

| Variable | Estimate | Standard error | p-value | Odds ratio |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| Intercept | –3.786 | 0.732 | 0.000 | 0.02 |
| international plan (q) | 1.205 | 0.418 | 0.004 | 3.34 |
| voice mail plan (q) | –1.123 | 0.191 | 0.000 | 0.33 |
| total day minutes (c) | 0.005 | 0.002 | 0.004 | 1.01 |
| total eve minutes (c) | 0.007 | 0.001 | 0.000 | 1.01 |
| total night minutes (c) | 0.006 | 0.001 | 0.000 | 1.01 |
| total intl calls (c) | –0.073 | 0.030 | 0.014 | 0.93 |
| total intl charge (c) | 0.417 | 0.099 | 0.000 | 1.52 |
| state NJ (q) | 0.830 | 0.402 | 0.039 | 2.29 |
| state RI (q) | –1.530 | 0.733 | 0.037 | 0.22 |
| state MT (q) | 1.293 | 0.482 | 0.007 | 3.64 |

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| state VA (q) | –1.393 | 0.681 | 0.041 | 0.25 |
| state AZ (q) | –1.285 | 0.612 | 0.036 | 0.28 |
| state WY (q) | –1.071 | 0.538 | 0.046 | 0.34 |
| state IL (q) | –3.150 | 1.121 | 0.005 | 0.04 |
| state CA (q) | 1.664 | 0.693 | 0.016 | 5.28 |
| leaf ID 5 (q) | 0.939 | 0.327 | 0.004 | 2.56 |
| leaf ID 6 (q) | –0.893 | 0.475 | 0.060 | 0.41 |
| leaf ID 7 (q) | –2.664 | 0.261 | 0.000 | 0.07 |

Source: own elaboration by using STATISTICA 10.0.

In the hybrid approach there are several predictors that significantly contributed to the model and have high positive effects with respect to the response (increase the probability of churn):

– "international plan" (q) – the probability of churning among customers with the international calling plan is almost 3.5 times higher than the probability among customers without that plan, one can notice a decrease in the value of the odds ratio in comparison to the basic logistic model,
– "total intl charge" (c) – for every unit increase in "total intl charge" the odds of churning increase by approximately 52%,
– "state NJ" (q) – clients living in the state of New Jersey have approximately 2 times greater odds of churning than clients from the states not included in the model (increase in the odds ratio from 2.17 in the basic model to 2.29 in the hybrid model),
– "state MT" (q) – customers from Montana have 3.64 times higher probability of churning than customers from the states not included in the model (the odds ratio value in the basic logistic model is equal to 3.65),
– "state CA" (q) – customers living in California have about 5 times greater odds of churning than customers from the states not included in the model (decrease in the odds ratio from 6.09 to 5.28),
– leaf ID 5 (q) – the probability of churning among the clients from the terminal node ID 5 (those for whom the daily number of minutes is lower than or equal to 246.6 and the number of calls to call center exceeds 3) is about 156% higher than the probability among customers who talk longer (reference category – leaf ID 3).

Terminal nodes that significantly contributed to the model (they are highlighted with a shade of grey) adjusted other predictors from the basic logistic model. It turned out that the independent variable "number customer service calls" did not contribute to the hybrid model. On the other hand it enriched the interpretation of the model by detecting interactions of predictors. In the basic logistic model (the main effects model) the relationship between "total day minutes" and churning was almost unobservable (the odds ratio was equal to 1.01).

### 3.5. Comparison of models' performance

To evaluate the models' performance several popular measures such as accuracy, recall, precision, true negative rate (TNR), G-mean and F-measure were used. The next three tables (Tables 4–6) show the misclassification matrices for the three models implemented in the test sample.

Table 4. Misclassification matrix for C&RT model

| Observed classes | Predicted classes | | Total |
|---|---|---|---|
| | no churn (0) | churn (1) | |
| No churn (0) | 1098 | 207 | 1305 |
| Churn (1) | 39 | 162 | 201 |
| Total | 1137 | 369 | 1506 |

Source: own calculations.

Table 5. Misclassification matrix for logistic model

| Observed classes | Predicted classes | | Total |
|---|---|---|---|
| | no churn (0) | churn (1) | |
| No churn (0) | 603 | 702 | 1305 |
| Churn (1) | 17 | 184 | 201 |
| Total | 620 | 886 | 1506 |

Source: own calculations.

Table 6. Misclassification matrix for hybrid C&RT-logit model

| Observed classes | Predicted classes | | Total |
|---|---|---|---|
| | no churn (0) | churn (1) | |
| No churn (0) | 863 | 442 | 1305 |
| Churn (1) | 27 | 174 | 201 |
| Total | 890 | 616 | 1506 |

Source: own calculations.

On the basis of the results from the misclassification matrices six measures were calculated: accuracy ((TP + TN)/(TP + FP + TN + FN)), recall (TP/(TP + FN)), precision (TP/(TP + FP)), true negative rate (TN/(FP + TN)), G-mean ((true negative rate × recall)$^{1/2}$) and F-measure ((2 × precision × recall)/(precision + recall)) – see details in Table 7. The best results were highlighted with a shade of grey. As one can easily see, the C&RT model outperforms other methods (except for recall). In general, the hybrid C&RT-logit model turned out to be better than the basic logistic model.

Table 7. Performance of models

| Measure | C&RT | Logistic model | Hybrid C&RT-logit |
|---|---|---|---|
| Accuracy | 0.837 | 0.523 | 0.689 |
| Recall | 0.806 | 0.915 | 0.866 |
| Precision | 0.439 | 0.208 | 0.282 |
| TNR | 0.841 | 0.462 | 0.661 |
| G-mean | 0.823 | 0.650 | 0.757 |
| F measure | 0.568 | 0.339 | 0.426 |

Source: own calculations.

The values of the performance measures are confirmed by the cumulative lift chart (Figure 3) and the cumulative gain chart (Figure 4). One can see the higher predictive power of the decision tree model in comparison with the hybrid approach and the basic logistic model. This means in this case that hybridization allowed only to enrich the interpretation of the model and to detect quickly the interaction between the variables. The researcher can interpret the odds ratios which are not directly available in the C&RT model, however losing the predictive properties of the model. It should be noted that in terms of the lift measure, the difference between the models becomes smaller and smaller starting from the third decile. In the third, the fourth and the fifth decile cumulative lift measures for the C&RT model are as follows: 2.95, 2.32 and 1.86, while for the hybrid model these values are respectively: 2.80, 2.16 and 1.81.

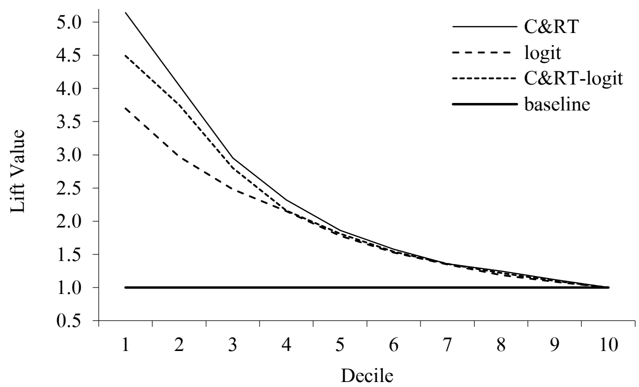


Fig. 3. Cumulative lift chart

Source: own calculations.

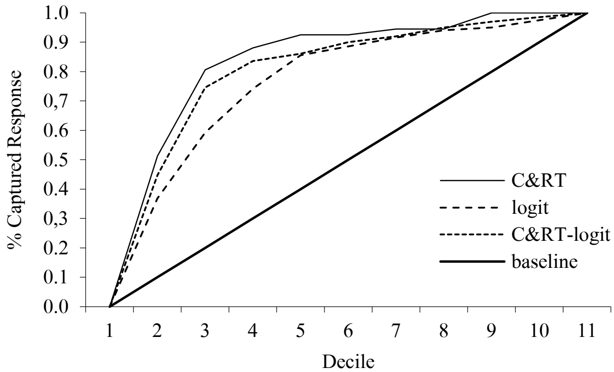


Fig. 4. Cumulative gains chart

Source: own calculations.

## Conclusions

Prior to the building of the hybrid C&RT-logit model one counted on benefits from combining advantages of both analytical tools. It turned out that the interactions detected by the decision tree significantly contributed to the logistic model and enriched its interpretation. From the point of view of performance measures, the hybrid approach delivered better results than the basic logistic model, however, it was outperformed by decision tree. As far as the lift measure is concerned, differences between C&RT and C&RT-logit are relatively small starting from the third decile. The hybrid model based on this particular dataset has a higher recall rate when compared to the single decision tree. Through odds ratios the interpretation of the model can be extended beyond the "if… then…" rules. An important advantage of this approach is also obtaining unique predicted probabilities for each case from the test sample. The remaining results did not fully meet the expectations of the author, however, the analysis of other datasets leads to optimistic conclusions, because there are situations in which the hybrid C&RT-logit model outperforms single decision tree[21]. In general, the researcher should decide whether he focuses on a higher predictive power of the model or agrees to a lower performance in return gaining an expanded interpretation of the model and the probabilities of churning assigned to individual cases. The experiment certainly should be extended to other datasets with binary dependent variable relating to analytical CRM.

## Notes

1 The project was financed by a grant from National Science Centre (DEC – 2011/01/B/HS4/04758).

2 Gupta, Zeithaml (2006), p. 721.

3 Gummesson (2008), p. 105.

4 Ibidem, pp. 280–281.

5 Sheth, Parvatiyar (2000), p. 191.

6 Ibidem, p. 194.

7 *Customer Retention...* (2004), p. 2.

8 Christopher et al. (2008), p. 8.

9 Jeffery (2010), p. 92.

10 Farris et al. (2006), p. 135

11 Phadke (2013); Idris et al. (2013); Liao, Chueh (2011).

12 Van den Poel, Larivière (2004); Kim et al. (2005); Naveen et al. (2010).

13 Miguéis et al. (2012); Buckinx, Van den Poel (2005).

14 Morik, Köpcke (2004).

15 Burez, Van den Poel (2007).

16 Coussement, Van den Poel (2008).

17 Giudici, Dequarti (2011).

18 Breiman et al. (1984).

19 Lindahl, Winship (1994).

20 Steinberg, Cardell (1998).

21 The results of the hybrid C&RT-logit models based on different datasets were presented during the European Conference on Data Analysis (ECDA 2014) in Bremen (M. Łapczyński, *The use of hybrid predictive C&RT-logit models in analytical CRM*, Second European Conference on Data Analysis, ECDA 2014, July 2–4, Jacobs University, Bremen, “Program and Abstracts”, p. 42).

## References

Breiman, L., Friedman, J., Stone, C.J. & Olshen, R.A. (1984). *Classification and Regression Trees*. New York: Chapman and Hall.

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, 252–268.

Burez, J. & Van den Poel, D. (2007). CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. *Expert Systems with Applications*, 32, 277–288.

Christopher, M., Payne, A. & Ballantyne, D. (2008). *Relationship Marketing. Creating Stakeholder Value*, 5 ed. Oxford: Butterworth-Heinemann.

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

*Customer Retention: A Business Imperative* (2004). Baxter Strategies Incorporated. New York: Woodbury.

Farris, P.W., Bendle, N.T., Pfeifer, P.E. & Reibstein, D.J. (2006). *Marketing Metrics: 50+ Metrics Every Executive Should Master*. New Jersey: Wharton School Publishing.

Giudici, P. & Dequarti, E. (2011). Statistical Models to Predict Academic Churn Risk. In: *Classification and Multivariate Analysis for Complex Data Structures*, eds. B. Fichet, D. Piccolo, R. Verde, M. Vichi (pp. 41–49). Berlin, Heidelberg: Springer-Verlag.

Gummesson, E. (2008). *Total relationship marketing*. 3 ed. Oxford: Butterworth-Heinemann.

Gupta, S. & Zeithaml, V. (2006) Customer Metrics and Their Impact on Financial Performance. *Marketing Science*, 29 (6), 718–739.

Idris, A., Khan, A. & Lee, Y.S. (2013). Intelligent churn prediction in telecom: employing mRMR feature selection and RotBoost based ensemble classification. *Applied Intelligence*, 39, 659–672.

Jeffery, M. (2010). *Data-Driven Marketing. The 15 Metrics Everyone in Marketing Should Know*. Hoboken, NJ: John Wiley & Sons.

Kim, S., Shin, K.S. & Park, K. (2005). An Application of SVM for Customer Churn Analysis: Credit Card Case. In: *Advances in Natural Computation. Lecture Notes in Computer Science*, eds. L. Wang, K. Chen, Y.S. Ong (pp. 636–647). Berlin, Heidelberg: Springer-Verlag. Vol. 3611.

Liao, K.H. & Chueh, H.E. (2011). Applying Fuzzy Data Mining to Telecom Churn Management. In: *Intelligent Computing and Information Science. Communications in Computer and Information Science*, ed. R. Chen (pp. 259–264). Berlin, Heidelberg: Springer-Verlag. Vol. 134.

Lindahl, W.E. & Winship, C. (1994). A Logit Model with Interactions for Predicting Major Gift Donors. *Research in Higher Education*, 35 (6), 729–743.

Łapczyński, M. (in press), *The use of hybrid predictive C&RT-logit models in analytical CRM*, Second European Conference on Data Analysis, ECDA 2014, July 2–4, Jacobs University, Bremen, “Program and Abstracts”, http://ecda2014.eu.

Miguéis, V.L., Van den Poel, D., Camanho, A.S. & Falcão e Cunha, J. (2012). Predicting partial customer churn using Markov for discrimination for modeling first purchase sequences. *Advances in Data Analysis and Classification*, 6, 337–353.

Morik, M. & Köpcke, H. (2004). Analysing Customer Churn in Insurance Data – A Case Study. In: *Proceedings of the 8th European Conference on Principles and Practice of Knowl-*

*edge Discovery in Databases*, eds. J.F. Boulicaut, F. Esposito, F. Giannotti, D. Pedreschi (pp. 325–336). Berlin, Heidelberg: Springer-Verlag.

Naveen, N., Ravi, V. & Rao, C.R. (2010). Data Mining via Rules Extracted from GMDH: An Application to Predict Churn in Bank Credit Cards. In: *Knowledge-Based and Intelligent Information and Engineering Systems. Lecture Notes in Computer Science*, eds. R. Setchi, I. Jordanov, R.J. Howlett, LC. Jain (pp. 80–89). Berlin, Heidelberg: Springer-Verlag. Vol. 6276.

Phadke, C., Uzunalioglu, H., Mendiratta, V.B., Kushnir, D. & Doran, D. (2013). Prediction of Subscriber Churn Using Social Network Analysis. *Bell Labs Technical Journal*, 17 (4), 63–76.

Sheth, J.N. & Parvatiyar, A. (2000). Relationship Marketing in Consumer Markets. Antecedents and Consequences. In: *Handbook of Relationship Marketing*, eds. J.N. Sheth, A. Parvatiyar (pp. 171–208). Thousand Oaks, CA: Sage Publications.

Steinberg, D. & Cardell, N.S. (1998): *The hybrid CART-logit model in classification and data mining*, www.salford-systems.com (14.10.2007).

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.