Közgazdász Fórum
Forum on Economics and Business
18 (123), 3–33.
2015/2

Credit portfolio analysis using the Churn model – A case study[1]

BÉLA-GERGELY RÁCZ[2] – ANCA BORZA[3] – CSILLA ÁBRÁN[4]

The main objective of this article is to present a credit portfolio analysis framework for small and medium-sized enterprises, which, if applied, may bring them significant competitive advantages. While several scientific results have been published on bank credit portfolio analysis in the past decades, the literature in the field rarely focuses on the SME sector, where there may be a significant customer risk. The proposed framework analyses customers' previous behaviours, and based on the available historical data, draws conclusions concerning their future behaviour. By using the Churn model it can be decided whether a customer falls into one or another customer category (is it a committed, enthusiast, satisfied customer, a customer building on habits or one that likes to switch frequently between suppliers). In the increasingly acute market competition, companies are losing windfall revenues and profits due to the fact that their customers switch to another provider. It is also not negligible that the cost of winning a new customer is significantly higher than that of retaining old customers. The results are relevant not only from a theoretical point of view, but also from a practical perspective, given that based on the analysis we shall build a decision-making instrument (credit portfolio analysis framework) which can be easily applied in practice.

Keywords: commercial credit portfolio, credit risk, Churn model, scoring model.

JEL code: G320.

[1] This work was co-financed from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013, project number POSDRU/187/1.5/S/155656 "Help for doctoral researchers in economic sciences in Romania".

[2] PhD student, teaching assistant, Babeş–Bolyai University, Faculty of Economics and Business Administration, e-mail: bela.racz@econ.ubbcluj.ro.

[3] PhD, professor, Babeş–Bolyai University, Faculty of Economics and Business Administration, e-mail: anca.borza@econ.ubbcluj.ro.

[4] MSc student, Babeş–Bolyai University, Faculty of Economics and Business Administration, e-mail: abrancsilla@gmail.com.

Published: Hungarian Economists' Society from Romania and Department of Economics and Business Administration in Hungarian Language at Babeș–Bolyai University
ISSN: 1582-1986 www.econ.ubbcluj.ro/kozgazdaszforum

Introduction

The increasingly acute market competition determines companies, particularly those in developed countries, to pursue customer-oriented marketing policies. In saturated markets, market participants may even apply such strategies by which companies seek to attract customers from other competitors. In such situations it is essential to create customer loyalty programmes, which prevent customer attrition.

The *Churn model* is a perfect method for companies to identify customers prone to churn, and to include in their strategy actions for making them loyal customers (Kim–Yoon 2004).

The aim of our research is to map and to analyse the customer portfolio of MarsoRom Ltd. – the chosen case study company from Romania – in order to create added value for the business. An additional goal of our research is to elaborate a method, which can be easily adapted and used by other companies, too. The available historical data allows us to analyse the behaviour of customers so far, and enable us to draw conclusions with regard to their future behaviour. By using the Churn model it can be decided whether a customer falls into one or another customer category (is it a committed, enthusiast, satisfied customer, a customer building on habits or one that likes to switch frequently between suppliers).

Companies are losing windfall revenues and profits also due to the fact that their customers switched to another provider. At the same time, it is not negligible that the cost of winning a new customer is significantly higher than that of retaining old customers. Long-standing customers are willing to pay a higher price if they are able to benefit from a high-quality service for a long period of time; they will have lower handling costs and profits will increase over time. Our first hypothesis is that the Churn model can be easily applied for small and medium- sized enterprises (SMEs), it is a useful credit scoring tool, which is capable to create value for the company. We have developed a scoring system measuring customers' bankruptcy probability, which has low implementing costs. According to Reicheld and Sasser (1990), companies could increase their profits by up to 100%, if their customer loyalty was 10% higher. Our second hypothesis is that if companies reduced the number of customers prone to attrition by 5%, their revenue could increase by at least 5%.

Finally, we elaborate a strategy which, based on the Churn model, will work on reducing by 5% the attrition of customers with a low bankruptcy probability.

The proposed analysis framework will be presented through this case study, but the method can be applied to any small and medium-sized enterprise concerning their own customers.

Literature review

Banks use complex scoring systems, and their operation is very difficult to understand correctly in the case of smaller banking institutions, often resulting in inappropriate implementations thereof (Han 2014). This is particularly true for the SME sector; therefore, it is required to elaborate a system that can successfully be applied in corporate practice.

As a result of intense competition in the industries, companies have to balance their net working capital as well as possible in order to be successful on the long run. Customer lending is one of its cornerstones, and if it happens irresponsibly, the company may become insolvent. If it provides fewer loans as recommended, it might lose customers (Lee et al. 2006).

Customers of products and services are well aware of the fact that commercial loans are always cheaper than bank loans (Giannetti et al. 2011), thus they always try to use up their credit limit as much as possible.

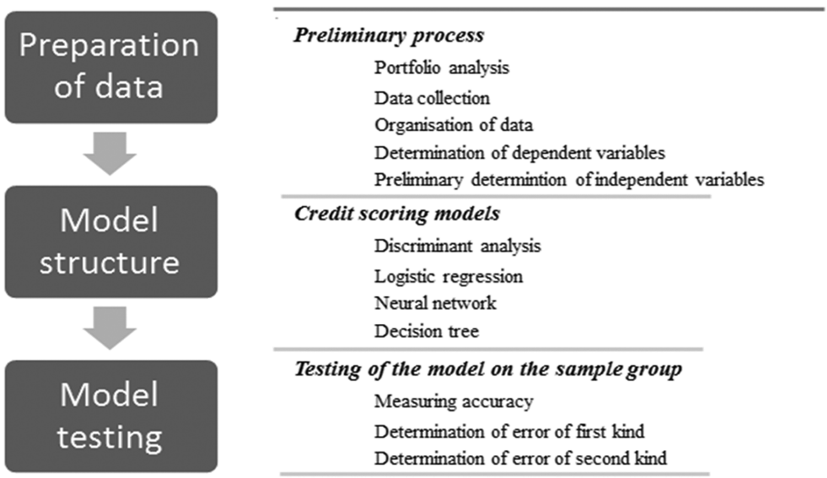
Credit scoring models

Credit scoring models are widely used especially by banks in order to measure credit risk (Mester 1997; Baesens et al. 2003). The aim of the models is to determine – based on the available customer data – whether a customer is creditworthy or not. Moreover, based on current available data, they try to establish the probability of a future non-payment.

The scoring model is the cornerstone of conscious lending. Today, keeping the credit portfolio at an acceptable quality level is unimaginable without an accurate and automated risk analysis system. The continuous validation and development of scoring models is an important aspect as it either contributes to lowering the portfolio risk or it leads to a bigger portfolio at an unaltered risk-taking level (Crook et al. 2007).

The model assigns a non-paying probability to each customer, which will be a characteristic thereof. The banks themselves decide the cut-off point under which they do not give out loans. The data used in the model may come from the system of credit institutions (the credit applicant is blacklisted, with previous non-payments, etc.) or may be personal customer data (such as age, income, gender, number of dependent persons, average monthly expenses, etc.)

A pre-requisite of the scoring model is the uploading of a database with data significantly influencing the non-paying probability of credit applicants. The dependent variable indicating the occurrence of bankruptcy or non-payment must be determined within the sample group. This is a dummy variable, which may get two values: 0 or 1. It is recommended to think of the independent variables, which may be significant concerning the model (Figure 1).



Source: Wealth Management System Limited 2014

Figure 1. The process of building a scoring model

The method we chose and applied, namely that of logistic regression, shall be described in detail later in this paper. Following the elaboration of a model it is essential to test it, in order to identify errors of first and second kind. The misclassification of good customers as bad (n_{FP}), i.e. their rejection indicates an error of first kind, while the classification of bad customers as good indicates an error of second kind (n_{FN}).

Figure 2 shows the classification matrix with the correct and erroneous customer classifications. The general form of the matrix is as follows:

		Predicted/forecast value		
		Good	**Bad**	
Actual value	**Good**	n_{TP}	n_{FP}	n_P
	Bad	n_{FN}	n_{TN}	n_N
		n_T	n_F	n

Source: Wealth Management System Limited 2014

Figure 2. Classification matrix

There are two ways to improve the given model: we either increase the rejection rate of bad customers, or decrease the rejection rate of good customers. The “n” variable measures the classification accuracy of the joint classification of good and bad customers.

A scoring system can be drawn up also in case of larger corporations, provided that sufficient data is available, which influences non-payment.

In order to better understand what a function forming up a scoring system looks like, let’s take a look at the Z-scoring model according to Altman, considered a benchmark scoring system by the literature (Kumar–Anand 2013):

$$Z = 1.2\ T_1 + 1.4\ T_2 + 3.3\ T_3 + 0.6\ T_4 + 0.999\ T_5$$

where

T1 = Working capital / Total assets – compares the amount of liquid assets to the size of the company;

T2 = Net profit / Total assets – measures the profitability of the company;

T3 = EBIT / Total assets – considers that the operating income as opposed to assets is important in terms of the long-term viability;

T4 = Equity / Debt – measures long-term payment capability;

T5 = Revenue / Total assets.

Table 1. Meaning of the scoring model according to Altman

Z score cut-off	Meaning of indicator	Observation
< 1.81	Bankruptcy	Danger zone, high probability of bankruptcy
1.81-2.99	Successful company	Difficult to estimate probability of bankruptcy
> 2.99	Stable company	Low probability of bankruptcy

Source: Kumar–Anand 2013

The Churn model

Companies usually make a distinction between voluntary churn and involuntary churn. Voluntary churn occurs due to a decision taken by the customer to switch to another company or service provider, involuntary churn occurs due to circumstances such as a customer's relocation to a long-term care facility, death, or the relocation to a distant location. In most applications, involuntary reasons for churn are excluded from the analytical models. Analysts tend to concentrate on voluntary churn, because it typically occurs due to factors of the company-customer relationship which companies control, such as how billing interactions are handled or how after-sales help is provided (Xin et al. 2009.14). Due to the unpredictability of involuntary churn, our study focuses on voluntary churn.

Reducing churn by 5% could increase future sales by 25-85% in a wide variety of sectors, such as insurance, banking, manufacturing or constructions (Ahmad–Buttle 2001), but this statement is true also for other industries. It is also much cheaper to retain an old customer than to attract a new one – this is one of the key concepts of marketing (Rosenberg–Czepiel 1984).

We distinguish between the following churn categories (Lazarov–Capota 2007):

• *Active:* the customer voluntarily switches to another provider. This may be due to the more favourable pricing policy of competitors, the customer's lack of satisfaction regarding the service provided, lack of information concerning the service, lack of customer loyalty rewarding, etc.

• *Accidental:* the customer relocates to another place, struggles with financial difficulties or becomes insolvent. Thus, the customer did not intend to switch providers.

• *Passive:* the customer goes bankrupt.

The essence of the Churn model is that the company, by analysing the behaviour of its customers, tries to define the circle of customers which will likely quit the service currently provided to them and switch to a competitor company. Identifying these customers and taking preventive measures is crucial in order to preserve the existing pool of customers. The already mentioned probability of non-payment is closely related to the Churn model, as the retention of customers does obviously not stand above all. Companies do not need to retain all of their customers. If the probability of non-payment is known, customised, preventive and cost-efficient measures and actions can be taken (Hadden et al. 2006).

The mission of the Churn model is thus to identify customers that are less loyal to the company and display an alternating behaviour. Making these customers loyal is a cheaper process than attracting new customers.

The application of the Churn model is widespread mainly among telecommunications companies, where customer churn is 30% per year. This results from the fact that the limits of switching to another provider are low. Nowadays the transaction cost switching to any other provider or supplier is very low, which brings about the intensification of competition and loyalty programmes offered by providers become important (Gropal–Meher 2008; Owczarczuk 2010).

Reinartz, Thomas and Kumar (2005) proved that if a company invests less in customer retention than in attracting new customers, this will have a much more negative impact of profitability. With the expansion of the press and of internet access, it is more convenient and simple for customers to switch between providers and suppliers.

It is essential to define the duration of churn for each sector and company. While in the case of telecommunications the time interval for a customer to switch to another provider may be somewhere between a couple of days and a couple of weeks, in the wholesale sector this interval may well be a couple of months or even a year.

There are certain factors which directly influence or identify the reasons of churn, such as time elapsed from becoming a customer (the initial period is the most critical), customer dissatisfaction, defective products, late delivery of products, etc. (Owczarczuk 2010).

It is worth elaborating within the Churn model the different customer categories prone to churn so that these bring the highest yield for the amount invested. There are certain customers in which it is not worth making extra investments, as no significant added benefits will be obtained due to the high probability of non-payment.

When elaborating the strategy for retaining customers prone to churn, the following considerations must be taken into account (Verbeke et al. 2012):

· when selecting customers with a high probability of churn it is important to examine their profitability;

· in order to develop a cost-effective strategy, the appropriate customer groups should be targeted;

· customers that will either way switch to another provider need to be ignored;

· loyal customers, which will not switch to other providers by any means shall be treated cost-effectively during the strategy to be adopted.

The concept of Churn models underwent many changes in the past 15–20 years, and current models use a wide range of different customer data in their forecast. Interestingly, this development cannot be observed in modelling algorithms, these are still mostly using the models used in the beginning, such as logistic regression, decision trees and neural networks (Tsai–Lu 2009).

Research methodology

For developing the Churn model and showing its benefits, we used a case study research. We identified a few companies, which fitted our research goals, and we approached them. We have chosen as the case study company the MarsoRom Ltd., because they were willing to cooperate in our current research. We have interviewed the management for a better understanding of the company's current situation. We also had access to all the financial data, which was a necessary criterion in building our statistical model.

The case study company is a tyre distributor company, which given its core business, also deals with commercial lending, as it offers its customers the possibility to pay at a later stage. Such debit appears as *accounts payable* in the balance sheet of the customer company, and as *accounts receivable* in that of the selling company. The control and management of receivables is essential for a company with such profile as outstanding receivables can significantly affect liquidity and long-term profitability. Therefore, it is important to analyse each customer's creditworthiness, and if a customer is deemed as such, the maximum credit limit must also be established. In the SME sector there are often insufficient funds for outsourcing this task, and companies usually prefer to do this internally. With the use of historical data an analysis of the past behaviour of customers can be performed and conclusions can be drawn with regards the future. By using the Churn model it can be decided whether a customer falls into one or another customer category (is it a committed, enthusiast, satisfied customer, a customer building on habits or one that likes to switch frequently between suppliers). By using this model, a strategy can be developed for the prevention of customer attrition, and thus for the improvement of the company's efficiency. One of the issues raised is the control and management of accounts receivable, which mainly cause problems when they are no longer collectable. In such case it is essential to introduce a scoring system which facilitates the answering to the questions concerning the possible increase of the credit limit provided to existing customers, as well as helps to determine the creditworthiness of new customers. As it is revealed by our analysis, the churn of existing customers results in a massive loss of revenue, and acquiring new customers is costlier than retaining existing customers. In the case of MarsoRom the churn rate of wholesale customers is 25.3%, which, although is less than that observed in the case of telecommunication companies, is outstanding in relation to the sector (the average churn in the sector is 15%).

Our goal is to draw up a list of customers potentially prone to churn using the methods presented above, and develop a strategy for the company analysed in the case study, which could decrease by 5% the churn rate. Afterwards we will determine the impact of the strategy on the profitability of the company and see whether it is worth adapting future strategies to this perspective.

Multivariable regression analysis

The characteristics of bivariate regression can be simply generalised to the simultaneous analysis of three or more variables. In such case the equation looks like this:

$$y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_N X_N,$$

where

X_i – different regression variables;

a_i – regression coefficients.

Among independent variables not only continuous variables are permitted, but nominal (dummy) variables as well. The multivariate analysis, i.e. involving several variables in the analysis makes it more valuable and complex, as the data to be obtained is more versatile. However, it must be taken into consideration that in case of multiple variables the result is more difficult to interpret.

In the case of a multivariate analysis one of the most important factors is that X_i variables should be independent from each other, i.e. there should be no relation between variables. *It is called multicollinearity.* Variables related to each other should be excluded from the analysis. For the analysis of multicollinearity the determinant of the correlational matrix of variants can be used: in case of $R = 0$, the relation between variables is maximal, and for $R = 1$ the variables are independent (Fidy–Makara 2005).

The outlier as well as the limited number of cases may strongly influence the result of the analysis. In case of a multivariate analysis the rule of thumb regarding the number of cases is: the required number of cases should be at least six times higher than the number of X variables.

Logistic regression

The discriminant analysis bears a resemblance with the regression analysis in that it tries to express a dependent variable with the help of independent variants. In the case of the discriminant analysis we use categorical dependent variables besides continuous dependent variables. One of its disadvantages is also that it requires linearity, normality and preliminary group probabilities (Oravecz 2009).

As opposed to the discriminant analysis, the *logistic regression* dissolves the condition of the normal distribution of independent variables, and it is able to use qualitative criteria, and with it the weight of independent variables can also be obtained.

The basis of logistic regression is odds, which in the case of scoring models covers the:

$$odds_x = \frac{P_{survivor/x}}{1 - P_{survivor/x}}$$

quotient of the occurrence probability of survival and bankruptcy. The logit method uses the supposition that the logarithm of the odds ratio linearly depends on the explanatory variables:

X_j – independent variables;

β_j – regression parameters.

Binary logistic regression groups individuals to the defined area. This model is useful when the variable is dichotomous (e.g. either buys or not, either pays or is insolvent). Binary logistic regression can also be used for forecasting the probability of certain events.

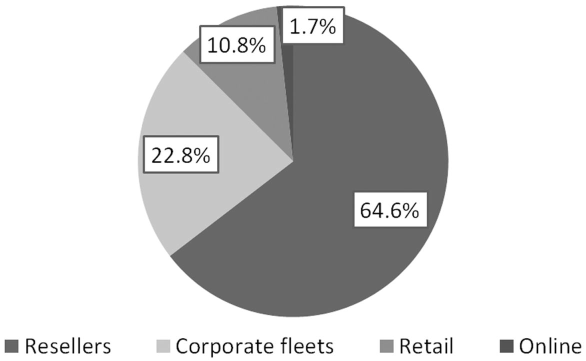
$$\text{logit} = \ln(\text{odds}) = \beta^T x$$

$$P(\text{solvent}) = \frac{e^Z}{1 + e^Z} = \frac{e^{\beta_0 + \Sigma(\beta_j x_j)}}{1 + e^{\beta_0 + \Sigma(\beta_j x_j)}}$$

The disadvantage of the logistic regression is that it reacts very sensitively to the correlations between independent variables, outliers or missing values and it is inefficient on small samples (Lázár 2011). Nowadays the logit model is the most widely used classification method in the field of credit scoring. The main reasons for this are: it is easily interpretable, it performs well, and not only does it classify but it also estimates the probability of defaulting (Basel II. rule). Furthermore, the method does not require the normal distribution of explanatory variables, thus categorical dependent variables can be easily integrated in it (Oravecz 2009).

Presentation of data

In order to identify what events or financial indicators lead to bankruptcy or financial collapse we studied the 2013 client portfolio of the firm. The portfolio consists of resellers, corporate fleets, online sale and retail. There is no financial risk in selling in retail or online due to the cash payment, on the other hand the risk rises with granting credit limits to corporate clients and resellers.



Source: own design based on company data

Figure 3. Composition of MarsoRom's client portfolio in 2013 (%)

In 2013 the uncollectible receivables were five times bigger than in 2012. This statistic brought up the question of what leads to the insolvency of resellers and corporate clients. Table 2 enlists the statistical data used by the default prediction and Churn models.

Table 2. Data used in the scoring model

Data	Explanation
Default	The client went bankrupt or did not pay its debt
Churn	The client chose another supplier within a year
Credit limit	The amount of credit that the company extends to a client
100 day debt	The client hasn't paid its obligation for more than 100 days
Payment method	The way the client will pay for the product (ex. cash payment, check payment, bank transfer)
Client category	MarsoRom's client portfolio consists of the following categories: retail, resellers, online and corporate fleets
Quantity	Purchased quantity (RON)
Fixed assets	Fixed assets of the client from its balance sheet
Current assets	Current assets of the client from its balance sheet
Shareholder's equity	Shareholder's equity of the client from its balance sheet
Revenue	Client's revenue from its profit-loss statement

Source: own editing based on the data made available by the company and by the Ministry of Finances

In the case of the Churn model the dependent variable is the *churn* variable. When its value equals 1, it means that the companies that were MarsoRom's customers in 2012 were no longer its customers in 2013.

Furthermore, in the database we can also find the *credit limit* provided to customers. If a given customer guarantees repayment with a blank check, or has been the company's customer for a long period of time, then the company can purchase the products with delayed payment up to a given limit.

Accounts receivable over 101 days overdue are probably never going to be paid, as these usually characterise companies going through a difficult financial situation. Of course there are exceptions, as some customers do pay, but with a delay.

Payment terms indicate whether the given customer offers any guarantee for paying the debt. When developing the Churn model it may be relevant which *agent* was dealing with this customer, as the agreement concluding policy of agents may influence customer churn.

Besides this we also provided the value of the customer's purchases from MarsoRom.

We also included in the table the balance sheet data made available by the Ministry of Finances. The table contains only those data, which are relevant from the models' point of view.

Presentation of the model

In order to have a clear insight of the default mechanism of certain clients, we have examined the correlation between the default of a client and other balance sheet or MarsoRom data. We can notice medium correlation between default and Credit limit/Sales/Payment delay, and low correlation between default and Revenue/Liabilities and Equity. The higher the credit default category of the client, the higher its credit risk is (Figure 4).

```
Correlation Coefficients, using the observations 1 - 571
(missing values were skipped)

5% critical value (two-tailed) = 0,0823 for n = 568

        Default Client_category   credit_limit          Sales
         1,0000          0,5732        -0,4088        -0,3126 Default
                         1,0000        -0,3993        -0,2691 Client_category
                                        1,0000         0,7868 credit_limit
                                                       1,0000 Sales

        Revenue     Liabilities         Equity  Payment_delay
        -0,1129         -0,0599        -0,0537         0,3101 Default
        -0,1422          0,0005        -0,1004         0,1284 Client_category
         0,2434          0,1926         0,1645        -0,1359 credit_limit
         0,1485          0,1360         0,0790        -0,1081 Sales
         1,0000          0,6299         0,7624        -0,0343 Revenue
                         1,0000         0,1381        -0,0151 Liabilities
                                        1,0000        -0,0156 Equity
                                                       1,0000 Payment_delay
```

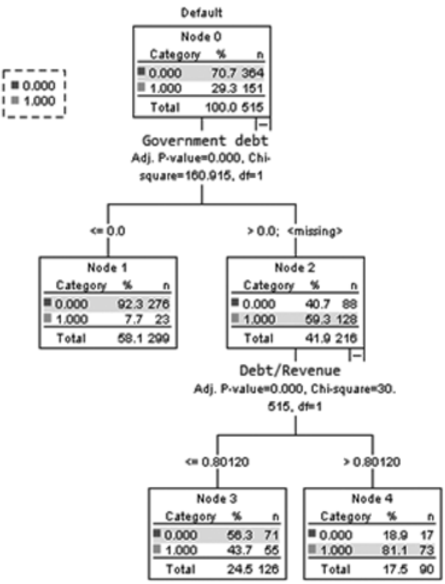
Source: company data analysed with Gretl software

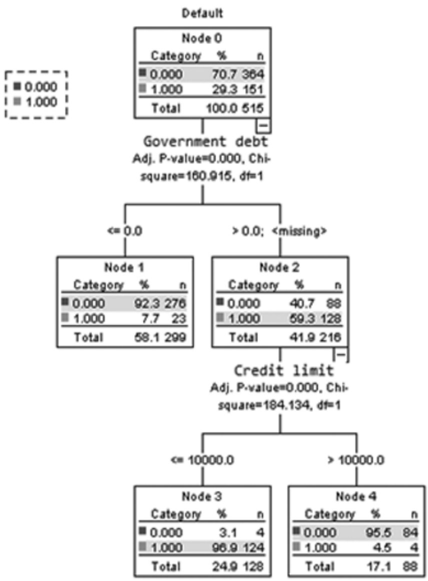
Figure 4. Correlation matrix

With the help of decision trees (Quinlan 1986), we have distinguished the clients by certain balance sheet characteristics. Figure 5 shows a significantly higher credit default risk among those clients, who failed to pay their taxes in time. Clients with delayed tax payment and debt/revenue ratio above 0.8 had an 81.1% default rate within the sample group. In case of clients with delayed tax payment and a debt/

revenue ratio lower than 0.8 the default rate is 43.7%. This means that a higher indebtedness leads to greater default probability.

Figure 6 describes the counterparty risk of clients from the tax indebtedness' and set credit limit's perspective. Those clients with delayed tax payments and low credit limit have a high risk of default (96.9% within the sample group). In case of clients with a higher credit limit than 10 000 RON, the default probability is significantly lower (4.5% in case of the sample group). This means that high credit exposure is limited to smaller credit limits within MarsoRom.





Source: company data analysed with SPSS software

Figure 5. Decision tree No. 1 **Figure 6. Decision tree No. 2**

Tables 3 and 4 show us the classification matrix of the two decision tree models. The first decision tree model could identify correctly 81.6% of the defaulted cases, while the second model had a 94% overall accuracy.

Table 3. Classification matrix (Decision Tree 1)

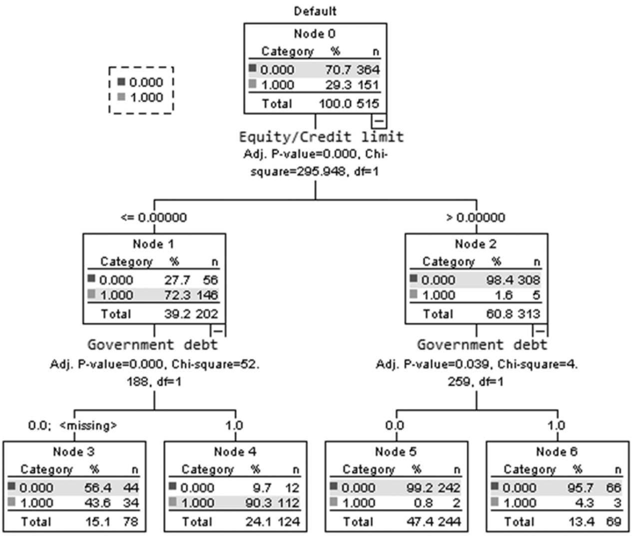
Observed	Predicted 0	Predicted 1	Percent Correct
0	347	17	95.3%
1	78	73	48.3%
Overall Percentage	82.5%	17.5%	81.6%

Table 4. Classification matrix (Decision Tree 2)

Observed	Predicted 0	Predicted 1	Percent Correct
0	360	4	98.9%
1	27	124	82.1%
Overall Percentage	75.1%	24.9%	94.0%

Source: company data analysed with SPSS software

Figure 7 shows how important the level of capital in terms of credit risk is. Those companies that are "neglected" by their owners are bound to default. Constant reinvestment of profit usually leads to a healthy functionality. In Romania negative capital does exist and shows disastrous results regarding counterparty risk (clients with negative private equity and accumulated government debts, had a 90.3% default rate within the sample group). This 3rd decision tree could identify with 90.3% accuracy the defaulting clients within the sample group (Table 5).



Source: company data analysed with SPSS software

Figure 7. Decision tree No. 3

Table 5. Classification matrix (Decision Tree 3)

Observed	Predicted 0	Predicted 1	Predicted Percent Correct
0	352	12	96.7%
1	39	112	74.2%
Overall Percentage	75.9%	24.1%	90.1%

Source: company data analysed with SPSS software

Based on the above statistics, MarsoRom has elaborated a client categorisation system that has been embedded into the client relations policy of the company. Clients from category A are usually multinational or national companies that have satisfying capital and over 5 million RON revenue. Category B consists of those companies that are significantly smaller than their multinational counterparts and have lower than 60% debt per revenue ratio. Clients with ratings C or D have a higher credit risk exposure due to the increased debt/revenue ratio, low profit margin and low capitalization.

According to MarsoRom's client relations policy, companies rated E can only be served with cash payment; this is due to unsatisfactory capitalization and a high debt per revenue ratio (Table 6).

Table 6. Characteristics of the credit risk based client categories

Client categories	Annual Revenue	Private equity	Debt/Revenue ratio	Profit/Revenue ratio
A	> 5 000 000 RON	> 1 000 000 RON	< 40%	> 5%
B	500 000 – 5 000 000 RON	45 000 – 1 000 000 RON	< 60%	1% – 5%
C	> 100 000 RON	> 15 000 RON	60% – 80%	> 1%
D		15 000 RON – 50 000 RON	80% – 150%	
E		< 50 000 RON	>150%	< 1 %

Source: MarsoRom's client relations policy

We have managed to build up a scoring model based on client category, granted credit limit and overdue indebtedness. Using the *Ordinary Least Squares* (OLS) method, we have constructed the model presented in Figure 8. All independent variables are significantly influencing the dependent variable, although we can notice heteroskedasticity and the lack of normal distribution.

```
Model 1: OLS, using observations 1-571
Dependent variable: Default

                    coefficient    std. error    t-ratio    p-value
  ---------------------------------------------------------------
  const            -0,124320       0,0435855     -2,852     0,0045     ***
  Clientcategory    0,160756       0,0120648     13,32      1,97e-035  ***
  Creditlimit      -9,02315e-07    1,65732e-07   -5,444     7,75e-08   ***
  Overduedebt       4,20174e-06    6,09741e-07    6,891     1,48e-011  ***

Mean dependent var    0,325744   S.D. dependent var    0,469063
Sum squared resid     73,24485   S.E. of regression    0,359416
R-squared             0,415964   Adjusted R-squared    0,412874
F(3, 567)             134,6102   P-value(F)            7,51e-66
Log-likelihood       -223,9165   Akaike criterion      455,8329
Schwarz criterion     473,2225   Hannan-Quinn          462,6173

White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 179,785
  with p-value = P(Chi-square(9) > 179,785) = 5,61848e-034

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 17,6708
  with p-value = 0,000145491
```

Source: company data analysed with Gretl software

Figure 8: Scoring model built using Ordinary Least Squares

In order to eliminate the heteroskedasticity problem we have elaborated two models, one based on the *Weighted Least Squares* (WLS) and one model using *Heteroskedasticity-corrected*. Both models are existing ones (P-value is smaller than 0.05). The adjusted R-square for the WLS model is 39.9%, while for the Heteroskedasticity-corrected model is 53.5%. This means that the model can correctly explain 53.5% of the new observations (Figures 9 and 10).

In both cases the client category and the overdue debt shows linear relationship with credit default, while the credit limit is inversely proportional. This proves that greater credit limits are only granted to financially stable companies.

Using binary logistic we managed to eliminate the problem of non-linearity and we have built a model that could identify with 96.1% precision the defaulting counterparts (Figure 10).

```
Model 2: WLS, using observations 1-571
Dependent variable: Default
Variable used as weight: Clientcategory

                  coefficient    std. error    t-ratio    p-value
  ----------------------------------------------------------------
  const           -0,146371      0,0584275     -2,505     0,0125     **
  Clientcategory   0,174156      0,0143264     12,16      2,23e-030 ***
  Creditlimit     -1,41112e-06   2,17718e-07   -6,481     1,98e-010 ***
  Overduedebt      3,48305e-06   5,88533e-07    5,918     5,64e-09  ***

Statistics based on the weighted data:

Sum squared resid    252,3647   S.E. of regression   0,667149
R-squared            0,401924   Adjusted R-squared   0,398759
F(3, 567)            127,0132   P-value(F)           6,21e-63
Log-likelihood      -577,0992   Akaike criterion     1162,198
Schwarz criterion    1179,588   Hannan-Quinn         1168,983

Statistics based on the original data:

Mean dependent var   0,325744   S.D. dependent var   0,469063
Sum squared resid    75,46943   S.E. of regression   0,364833

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 9,61539
  with p-value = 0,00816665
```

Source: company data analysed with Gretl software

Figure 9. Weighted Least Squares model

```
Model 3: Heteroskedasticity-corrected, using observations 1-571
Dependent variable: Default

                  coefficient    std. error    t-ratio    p-value
  ----------------------------------------------------------------
  const           -0,144096      0,0136622     -10,55     7,20e-024 ***
  Clientcategory   0,153299      0,00807340     18,99     1,35e-062 ***
  Creditlimit     -1,17115e-07   5,47442e-08   -2,139     0,0328    **
  Overduedebt      5,55038e-06   5,00127e-07    11,10     4,94e-026 ***

Statistics based on the weighted data:

Sum squared resid    1034,358   S.E. of regression   1,350654
R-squared            0,537759   Adjusted R-squared   0,535313
F(3, 567)            219,8773   P-value(F)           1,37e-94
Log-likelihood      -979,8430   Akaike criterion     1967,686
Schwarz criterion    1985,076   Hannan-Quinn         1974,470

Statistics based on the original data:

Mean dependent var   0,325744   S.D. dependent var   0,469063
Sum squared resid    77,42552   S.E. of regression   0,369531

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 20,1929
  with p-value = 4,12247e-005
```

Source: company data analysed with Gretl software

Figure 10. Heteroskedasticity-corrected model

```
Model 4 : Logit, using observations 1-571
Dependent variable: Default
Standard errors based on Hessian

                 coefficient    std. error       z        slope
  ---------------------------------------------------------------
  const          -1,21387       0,632226      -1,920
  Clientcategory  0,590073      0,150918       3,910    0,000498103
  Creditlimit    -0,000141362   2,14548e-05   -6,589   -1,19329e-07
  Clientrevenue   3,26272e-06   6,76073e-06    0,4826   2,75418e-09
  Overduedebt     0,000109090   4,19496e-05    2,601    9,20871e-08

Mean dependent var   0,325744   S.D. dependent var   0,469063
McFadden R-squared   0,696476   Adjusted R-squared   0,682601
Log-likelihood      -109,3815   Akaike criterion     228,7629
Schwarz criterion    250,4999   Hannan-Quinn         237,2433

Number of cases 'correctly predicted' = 549 (96,1%)
f(beta'x) at mean of independent vars = 0,001
Likelihood ratio test: Chi-square(4) = 501,98 [0,0000]

           Predicted
              0     1
  Actual 0  368    17
         1    5   181

Excluding the constant, p-value was highest for variable 4 (Clientrevenue)
```

Source: company data analysed with Gretl software

Figure 11. Scoring model built using binary logistic (1)

This 4th model showed no sign of collinearity problem (Figure 12).

```
Variance Inflation Factors

Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

Clientcategory    1,205
   Creditlimit    2,923
 Clientrevenue    2,642
   Overduedebt    1,026

VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient
between variable j and the other independent variables
```

Source: company data analysed with Gretl software

Figure 12. Scoring model built using binary logistic (2)

Results

As we mentioned it earlier, the customer churn for MarsoRom is 25.3%, which is considered high. The purpose of our analysis is to elaborate a strategy for bringing about a 5% decrease in customer churn.

This is important, as according to Ahmad and Buttle (2001), reducing customer churn by 5% can increase future sales by up to 80% in a wide range of sectors. Our second hypothesis is that by decreasing the customer churn of MarsoRom by 5%, a minimum of 5% increase in annual sales can be achieved.

In order to elaborate such a strategy, it is crucial to identify customers prone to churn. We applied logistic regression within the Churn model to identify customers likely to switch competitors. We applied the Churn model on the same sample group with a 99.3% processing rate (Table 7).

Table 7. Table summarising processed cases

Selected cases	**Included in the analysis**	1446	99.3
	Missing cases	10	0.7
	Total	1456	100.0
Number of cases not selected		0	0.0
Total		1456	100.0

Source: own calculations based on company data

Table 8 also helps in understanding the customer churn rate of the company, namely the above mentioned 25.3%. The company had 1456 customers in 2012, 366 of which either switched to another provider in 2013 or in some cases went bankrupt.

Table 8. Classification matrix

	Observed value		Forecast value		
			Churn		Correct classification
			0	1	
Step 0	Churn	0	1080	0	100.0%
		1	366	0	0.0%
	Cumulative probability				74.7%
The cut-off value is 0.3788					

Source: own calculations based on company data

By running the binary logistic regression presented in the section above, the model separated churn customers from loyal customers with a 72.4% probability. It classified correctly 234 customers as customers prone to churn and 813 as loyal customers. It calculated with 267 errors of first kind and 132 errors of second kind (Table 9).

Table 9. Classification matrix after running the regression

	Observed value		Forecast value		
			Churn		Correct classification
			0	1	
Step 1	Churn	0	813	267	75.3%
		1	132	234	63.9%
	Cumulative probability				72.4%
The cut-off value is 0.3788					

Source: own calculations based on company data

The cut-off value was set at 0.3788 in order to be able to identify customers prone to churn with a higher probability. The probability of bankruptcy is directly integrated in the Churn model, as it also significantly influences churn. Therefore, prior to developing the model, we need to apply for each customer the already presented scoring model, in order to find out their probability of bankruptcy.

Forecast_value = -0.4731 + 3.5458 *Default* + (-0.000187 *Credit limit*) + (-0.0000015 *Stocks*) + (-0.000005 *Quantity)* + (-0.000000004 *Accounts receivable*)

The credit limit influences customer churn by an inversely proportional manner. If a customer has a high credit limit at MarsoRom, the likelihood that it will switch to another provider is lower, and besides, it also risks not receiving a credit limit at the new provider. Consequently, the credit limit creates a barrier in front of customer churn.

The higher the purchasing amount of a customer from MarsoRom, the less probable the customer will switch to another provider. The larger the stocks of a customer, the less likely it will switch to another provider. However, the higher its accounts receivable, the more likely it will switch to a competitor company (Table 10).

Customers prone to bankruptcy are automatically regarded as prone to churn by the model, therefore it is essential to illustrate in the customer list whether or not that particular customer is prone to bankruptcy. According to the model, the company has a pool of almost 1000 loyal customers, which are not under bankruptcy proceedings.

Table 10. Independent variables in the equation

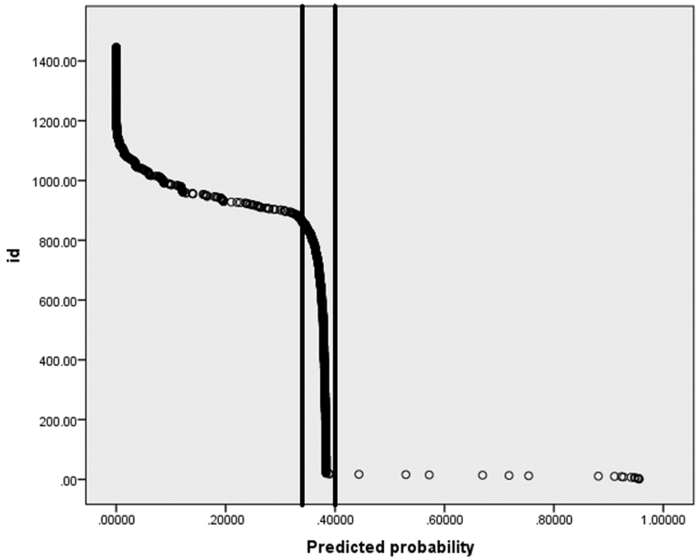
		B	S.E.	Sig.	Exp (B)
Step 1	Default	3.5458393999	0.997	0.000376	34.669
	Credit limit	0.0001870416	0.000	0.000000	1.000
	Quantity	0.0000050075	0.000	0.022664	1.000
	Accounts receivable	0.0000000443	0.000	0.002418	1.000
	Stocks	0.0000014891	0.000	0.002154	1.000
	Invariant	0.4731156095	0.072	0.000000	0.623

Source: own calculations based on company data

We have applied the developed Churn model to the 2013 customer portfolio of MarsoRom, and we obtained the distribution function presented in Figure 12. All cases located to the right of the vertical line are shown by the model as prone to churn. 100% of the cases from the sample group shown with a probability higher than 0.4% actually churned. The cases from the sample group located to the left of the lines churned with a probability of only 14%. Customers located between the two lines from the group of customers prone to churn.

Forming of customer groups prone to churn

In Table 11 we presented the churn groups created among the company's customer portfolio in 2014. In the first column we included the name of the group taking into consideration the estimated behaviour of the customers of the group. Customers prone to bankruptcy form a distinct category, which is associated with a 100% churn probability. The second column includes the churn probabilities predicted by the model, while the third column includes the probabilities of bankruptcy. In column 3 we included the bankruptcy probability observed in the given sample group in 2013. The last column shows the number of customers in the given group.



data

Figure 12. The distribution of the predicted probability of customer churn in the logistic regression

Table 11. Churn groups in the 2014 customer portfolio

Customer category	The groups formed in terms of the predicted churn values	Probability of bankruptcy in the 2014 customer portfolio	Probability of churn within the sample group (2013)	Number of customers in the 2014 customer portfolio
Prone to bankruptcy	0.4 – 1	88.5%	100%	26
Prone to churn	0.3838 – 0.4	0.0%	70%	63
Alternating	0.383 – 0.3838	0.0%	47%	386
Enthusiast	0.3788 – 0.383	0.0%	44%	815
Loyal customer	0 – 0.3788	0.2%	14%	1253

Source: own editing based on the Churn model

Consequently, we can affirm that in the 2014 customer portfolio of MarsoRom there are 1253 "loyal" customers with a churn probability less than 14%. The model identified 1201 customers which will likely switch to another supplier with a 50% probability. There are 63 customers with a higher churn probability of 70%; 0.1% of the customer portfolio is in danger of going bankrupt, but this group is irrelevant from the perspective of strategy development and it is important to differentiate these customers from those prone to churn but which are profitable.

The possible practical application of the above models: marketing strategy developed for different churn groups

By applying the above presented model to the company, the manager can become aware of the quality of its customers, and this information can also play a key role when taking strategic decisions.

Let's analyse the above customer groups according to the purchased amount. We can observe that customers identified as "enthusiasts" according to the Churn model have a 44% probability of switching to another provider. We further broke down this category according to the annual purchased amounts, obtaining a group which purchased products from MarsoRom in the 1000-2500 RON amount interval and another category with annual purchases between 2500 and 4500 RON. There are alternating customers which purchases between 700-1000 RON and there are customers prone to churn with purchases up to 700 RON (Table 12).

Table 12. The churn formed according to purchase value

Customer category	Churn probability	Number of customers	Annual purchase value (RON)	Recommended strategy	Drawing of lots
Prone to churn	70%	340	100 – 700	–	
Alternating	47%	190	700 – 1000	Recommendation I-II-III.	
Enthusiast1	44%	395	1000 – 2500	Recommendation I-II.	X
Enthusiast2	44%	229	2500 – 4500	Recommendation I.	X

Source: own editing based on the Churn model

In case of customers prone to churn we did not consider it was worth elaborating a separate marketing strategy, as compared to the average traded volume of the company, these customers purchase very small amounts and have a very high probability of churn, therefore a customer loyalty programme aimed at this group would be too costly. While MarsoRom receives a significant amount of marketing support from tyre manufacturers, including flyers, giveaways or other advertisement packages, the appropriate distribution of these marketing materials would generate significant windfall revenue, if they were forwarded to the appropriate loyal customers, which are also willing to purchase in larger volumes. Table 12 also includes which types of offers should be presented to which customer group. The strategy includes a gift programme based on a draw of lots, in which ten Formula 1 GOLD Tickets would be given away for the Hungarian Grand Prix. The customers eligible to participate in the draw are those who purchased products from MarsoRom for an amount above 3000 RON. Formula 1 gift would have a validity of 2 years and half of their cost would be borne by MarsoRom together with other manufacturers supporting the Hungarian Grand Prix, based on an agreement.

The offer packages mentioned in the strategy are included in Table 13. In order for the customer to benefit from the bonuses given after reaching the specified amount and by paying in time, it has to make a single purchase of the volume specified in the respective offer and pay in time.

Table 13. Parameters of the applied strategies

	Expected amount to be purchased (RON)	Bonus for the amount	Bonus for timely payment
Offer No. III	2 000	0.5%	0.05%
Offer No. II	3 000	0.6%	0.07%
Offer No. I	5 500	0.7%	0.08%
	10 000	0.9%	0.10%
	50 000	1.1%	0.12%
	100 000	1.3%	0.15%

Source: own editing based on the churn avoidance strategy developed

When calculating the benefits of the chosen strategy we estimated – based on historical company data – that 60% of the targeted customers would use any of the offers.

According to our calculations, by the end of 2014, MarsoRom will have 315 new loyal customers. This number meets the expected 5% decrease in customer churn. As the model only identified customers with relatively low purchases, the benefit of the strategy expressed in terms of revenue closely meets the expected 5% increase. Therefore, our hypothesis has been only partially confirmed.

If the churn avoidance strategy would be supplemented with a system for measuring the level of customer satisfaction towards the company's services, a 10% increase in revenues would be easily achievable, based on prior research on customer satisfaction (Taylor–Baker 1994). Thus, in the first year, the presented churn avoidance strategy is responsible for 4.7% of the revenue (Table 14).

Table 14. Breakdown of future expected revenue and expenditures, 2014–2016 (RON)

	2014	**2015**	**2016**
Total revenue	3 170 000	2 674 500	2 005 875
Gross profit (11%)	348 700	294 195	220 646
Total bonus for amounts	28 550	23 988	17 991
Total bonus for timely payment	2 402	1 345	1 008
Cost of Formula 1 tickets	6 000	6 000	0
Net profit	311 749	262 863	201 647
NPV net profit			**667 946**

Source: own editing based on the strategy developed

The present value of the cash-flow of the net profit of three years resulting from the churn avoidance strategy is 687 432 RON.

The practical development opportunities of the Churn model in the case of the company analysed

In order to identify with higher probability those customers that are prone to churn and which make high-volume purchases, it is worth keeping a record of customers' complaints or eventual remarks concerning the products and services of the company. By using a survey, customers prone to churn could be easier to identify. This obviously requires a high rate of processed data in order to bring results. Further independent variables directly influencing customers churn need to be looked for.

For the future, it is worth applying the Churn model together with other methods as well, such as the discriminant analysis. The database is also worth expanding with as much data as possible in order for the model to produce data as accurate and relevant as possible.

Conclusions and further research possibilities

In conclusion we can affirm that our first hypothesis is confirmed, because following and adapting the presented steps of the model, it can be easily, and almost cost freely implemented in a small and medium-sized company. For running the model, the company willing to use the presented method, should purchase the statistical analysis software, and should invest in human resource that will run it. The mentioned costs are significantly lower than the presented benefits of the model.

We also can affirm that the Churn model created an added value for the company, as it identified with almost 50% probability the customers characterised by an alternating behaviour or which are less loyal. Without the model and the churn avoidance strategy applied, the case study company would suffer annual revenue losses of up to 1.5 million RON. Due to the model and the related strategy, an added value of 687 432 RON is generated for MarsoRom, which clearly shows the benefits of applying the above presented model within the company.

While our second hypothesis is confirmed only partially, the difference between the expected and calculated growth of revenue is only 0.3% for the case study company, which could be lower in the case of other enterprises. The Churn model implies the application of the scoring model as well, as the probability of bankruptcy of companies also largely influences churn. In the scoring model we have only focused on the quantitative characteristics of a client. To further develop the model, qualitative information can also be crucial in the understanding of default probability. The quality of management, company strategies, client portfolio, product portfolio, suppliers and market position can also give us valuable data regarding default probability. By further improving the scoring model, the outstanding accounts receivable of the company could also be reduced. The analysed company deals with customer management and credit limit setting for customers on a daily basis. The credit scoring model qualifies each customer, thus it could facilitate daily customer evaluation processes and improve the accuracy of customer classification into the relevant non-paying customer categories. Competition is intense also in the Romanian tyre market, so for the future, the main priority for each company will be the retention of loyal and profitable customers. By elaborating such a model, MarsoRom could achieve a great deal of competitive advantage and strengthen its position on the Romanian tyre market.

Further development of the Churn model still has a great potential, as the company has several customers which purchase high volumes while having a high churn probability, although the model was unable to identify them. A method for the identification of these customers in the future could be carried out by the company by keeping a record of customer complaints and remarks and by conducting customer satisfaction surveys regarding the quality of the products and services provided by the company. The joint effects of different marketing strategies and Churn model outcomes could also be themes for future research. For a full proof of our hypotheses, further research should test our case study findings on a statistically relevant sample using questionnaires.

Bibliography

Ahmad, R.–Buttle, F. 2001. Customer retention: a potentially potent marketing management strategy. *Journal of Strategic Marketing* 9(1), 29–45.

Baesens, B.–Van Gestel, T.–Viaene, S.–Stepanova, M.–Suykens, J.–Vanthienen, J. 2003. Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society* 54(6), 627–635.

Crook, J. N.–Edelman, D. B.–Thomas, L. C. 2007. Recent developments in consumer credit risk assessment. *European Journal of Operational Research* 183(3), 1447–1465.

Fidy, J.–Makara, G. 2005. *Biostatisztika*. Budapest: InforMed.

Giannetti, M.–Burkart, M.–Ellingsen, T. 2011. What you sell is what you lend? Explaining trade credit contracts. *Review of Financial Studies* 24(4), 1261–1298.

Gropal, R. K.–Meher, S. K. 2008. Customer churn time prediction in mobile telecommunication industry using ordinal regression. In: Washio, T.–Suzuki, E.–Ting, K.–Ming, I. (eds.) *Advances in Knowledge Discovery and Data Mining*. Berlin: Springer, 884–889.

Hadden, A.–Tiwari, R.–Roy, D. 2006. Does Technology Matter? *International Journal of Intelligent Technology* 1(1), 104–110.

Han, C. 2014. Comparative analysis of credit risk models for loan portfolios. *Journal of risk model validation* 8(2), 3–22.

Kim, H. S.–Yoon, C. H. 2004. Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. *Telecommunications policy* 28(9), 751–765.

Kumar, M.–Anand, M. 2013. Assessing Financial Health of a Firm Using Altman's Original and Revised Z-Score Models: a Case of Kingfisher Airlines Ltd (India). *Journal of Applied Management and Investments* 2(1), 36–48.

Lázár, E. 2011. *Regressziós modellek alkalmazása a kereslet alapú árkutatásban.* PhD Thesis. Budapest: Corvinus University of Budapest.

Lazarov, V.–Capota, M. 2007. *Churn prediction.* http://www.vladislav.lazarov.pro/files/research/papers/churn-prediction.pdf, downloaded: 01.02.2015.

Lee, T. S.–Chiu, C. C.–Chou, Y. C.–Lu, C. J. 2006. Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis* 50(4), 1113–1130.

Mester, L. J. 1997. What's the point of credit scoring? *Federal Reserve Bank of Philadelphia Business Review* 3, 3–16.

Oravecz, B. 2009. *Szelekciós torzítás és csökkentése az adósminősítési modelleknél*. PhD Thesis. Budapest: Corvinus University of Budapest.

Owczarczuk, M. 2010. Churn models for prepaid customers in the cellular telecommunication industry using large data marts. *Expert Systems with Applications* 37(6), 4710–4712.

Quinlan, J. R. 1986. Induction of decision trees. *Machine Learning* 1(1), 81–106.

Reichheld, F. P.–Sasser, W. E. 1990. Zero defections: Quality comes to services. *Harvard Business Review* 68(5), 105–111.

Reinartz, W.–Thomas, J.–Kumar, V. 2005. Balancing acquisition and retention resources to maximize profitability. *Journal of Marketing* 69(1), 63–79.

Rosenberg, L. J.–Czepiel, J. A. 1984. A marketing approach for customer retention. *Journal of Consumer Marketing* 1(2), 45–51.

Taylor, S. A.–Baker, T. L. 1994. An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions. *Journal of Retailing* 70(2), 163–178.

Tsai, C. F.–Lu, Y. H. 2009. Customer churn prediction by hybrid neural networks. *Expert Systems with Applications* 36(10), 12547–12553.

Verbeke, W.–Dejaeger, K.–Martens, D.–Hur, J.–Baesens, B. 2012. New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research* 218(1), 211–229.

Wealth Management System Limited 2014. *Credit scoring model*. http://www.wealth.co.th/consult.php?consult_id=5, downloaded 01.02.2015.

Xin, Z.–Yi, W.–Hong-Wang, C. 2009. A mathematics model of customer churn based on PCA analysis. *Computational Intelligence and Software Engineering* 1, 1–8.