

Customer churn prediction using a hybrid method and censored data

Bahram Bahmani^{a*}, Golshan Mohammadi^a, Mehrdad Mohammadi^b and Reza Tavakkoli-Moghaddam^c

^aMasters of Financial Management, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

^bPhD Student, Department of Industrial Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran

^cDepartment of Industrial Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran

CHRONICLE

ABSTRACT

Article history:

Received January 22, 2013

Received in revised format

29 March 2013

Accepted 18 April 2013

Available online

April 20 2013

Keywords:

Customer churn prediction

Customer relationship
management

Neural network

Customers are believed to be the main part of any organization's assets and customer retention as well as customer churn management are important responsibilities of organizations. In today's competitive environment, organization must do their best to retain their existing customers since attracting new customers cost significantly more than taking care of existing ones. In this paper, we present a hybrid method based on neural network and Cox regression analysis where neural network is used for outlier data and Cox regression method is implemented for prediction of future events. The proposed model of this paper has been implemented on some data and the results are compared based on five criteria including prediction accuracy, errors' type I and II, root mean square error and mean absolute deviation. The preliminary results indicate that the proposed model of this paper performs better than alternative methods.

© 2013 Growing Science Ltd. All rights reserved

1. Introduction

Customers are believed to be precious assets in any organization and customer retention (CR) plays essential role on the success of any business unit in today's competitive environment. Customer relationship management (CRM) plans to understand and to measure the true value of customers and market segmentation is a technique for successful CRM (Judd et al., 1981; Berry & Linoff, 2003). Hung and Tsai (2008), for instance, concentrated on techniques, which provide a human manager with a visualized decision-making technique for market segmentation. They used a market segmentation technique, namely the hierarchical self-organizing segmentation model (HSOS), to deal with a real-world data set for market segmentation of multimedia on demand in Taiwan. The proposed HSOS was capable of providing a general idea of market segmentation step by step considered as an alternative method to other hierarchical cluster techniques for market segmentation.

Corresponding author. Tel: +989188729120
E-mail: mehrdadmohamadi@ut.ac.ir (M. Mohammadi)

Kim and Yoon (2004) implemented a binomial logit model using a survey of 973 mobile users in Korea and identified determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. The possibility that a subscriber would switch carrier depends on the level of satisfaction with alternative-specific service attributes including call quality, tariff level, etc. Nevertheless, only factors such as call quality, handset type, and brand image could influence customer loyalty as measured by the intention/non-intention to describe the service provider to various people. According to Gerpott et al. (2001) CR, customer loyalty (CL), and customer satisfaction (CS) are important objectives for telecommunication network operators. They used some information from a sample of 684 residential customers of digital cellular network operators in Germany and reported that all CR, CL, and CS must be treated as differential constructs, which are causally inter-linked. In their survey, mobile network operators' perceived customer care performance had no significant effect on CR and their findings recommended that an important lever for regulators to promote competition in cellular markets was the enforcement of efficient number portability procedures among mobile network operators.

Kim and Kwon (2003) investigated whether network size matters when new subscribers choose their service providers and reported that the intra-network call discounts and the quality signaling effect would likely to be the sources of the size effect. Madden et al. (1999) developed a model associated with the probability of subscriber churn to different service attributes and subscriber characteristics. Their results demonstrated that churn probability was positively associated with monthly ISP expenditure, but inversely associated with household income.

Van den Poel and Larivie're (2004) investigated the topic of customer attrition in the context of a European financial services firm and studied predictors of churn incidence as part of CRM. They contributed to the existing literature by combining various kinds of predictors into one comprehensive retention model including several "new" kinds of time-varying covariates associated with actual customer behavior. They also analyzed churn behavior based on a truly random sample of the total population using longitudinal data from a data warehouse. They reported that demographic characteristics, environmental changes and stimulating interactive and continuous' relationships with customers were of major concern when considering retention. In addition, customer behavior predictors only had a limited effect on attrition in terms of total products owned as well as the inter-purchase time.

Baesens et al. (2002) concentrated on purchase incidence modeling for a European direct mail firm using response models based on statistical and neural network techniques. They reported that Bayesian neural networks could offer a viable alternative for purchase incidence modeling. Coussement and Van den Poel (2008) used support vector machines in a newspaper subscription context to build a churn model with a higher predictive performance. They reported that only when the optimal parameter-selection procedure was applied, support vector machines could outperform traditional logistic regression, whereas random forests outperform both types of support vector machines. Hung et al. (2006) compared different data mining techniques, which could assign a 'propensity-to-churn' score periodically to each subscriber of a mobile operator. They reported that both decision tree and neural network techniques could deliver accurate churn prediction models by using customer demographics, billing information, contract/service status, call detail records, and service change log.

Keramati and Ardabili (2012) identified different factors, which would influence customer churn, the single most valuable of an organization's assets using one year's data from call log files relating to 3150 customers selected randomly from an Iranian mobile operator call-center database. They reported that a customer's dissatisfaction, their amount of service usage and certain demographic characteristics could impact on their decision to remain or churn. Ngai et al. (2009) in a survey

provided a roadmap to help future research and facilitated knowledge accumulation and creation concerning the application of data mining techniques in CRM.

According to Tsai and Lu (2009), churn management is a major task for companies to retain valuable customers, the ability to forecast customer churn seems to be necessary. Pendharkar (2009) implemented genetic algorithm based neural network techniques for predicting customer churn in cellular wireless network services. Seo et al. (2008) used a two-level model of customer retention in the US mobile telecommunication service market.

2. The proposed method

In this paper, we present a hybrid method based on neural network and Cox regression analysis where neural network is used for outlier data and Cox regression method is implemented for prediction of future events. The proposed Cox regression function has the following forms,

$$\log h_i(t) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}, \quad (1)$$

$$h_i(t) = \exp(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}), \quad (2)$$

In Eq. (2), $h_i(t)$ is the probability of customer churn at time t , x_{ij} is the j^{th} specification of customer i^{th} and β_i represents the coefficients to be estimated. Categorical specification is one of the advantages of Cox regression function. Note that there are different types of categories in mobile industry, which are categorized in terms of categorical functions such as age-range, busy/idle, etc. The proposed model of this paper uses a hybrid method described in Fig. 1 as follows,

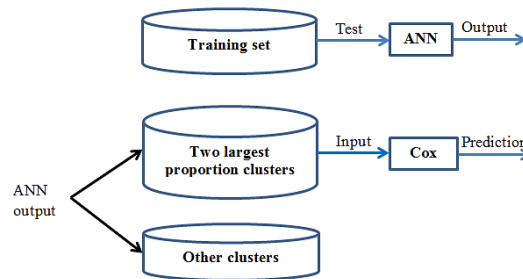


Fig. 1. The framework of the proposed method

The proposed model of the paper considers different factors influencing customer churn.

2.1. Customer dissatisfaction

The first group of hypothesis is as follows,

H_{1a} The number of disconnections calls positively influences customer churn.

H_{1b} The number of complains positively influences customer churn.

2.2. Level of service usage

The second group of hypothesis is as follows,

H_{2a} The monthly charge positively influences customer churn.

H_{2b} The call duration positively influences customer churn.

H_{2c} Too many digits of calling numbers positively influences customer churn.

H_{2d} The number of sms advertisements positively influences customer churn.

H_{2e} The number of unwanted calls positively influences customer churn.

2.3. Switching costs

There are some occasions where customers wish to switch to another operator and must pay some fine for doing this and this could increase customer churn. For the case study of this paper we only had one switching cost, which was associated with internet option. The following hypothesis is associated with this section,

H_3 The existence of internet option positively influences customer churn.

2.4. Demographic information

The other information, which may influence customer churn is associated with their customers' personal characteristics such as age. Table 1 demonstrates personal characteristics of the customers whose information are used for the proposed study of this paper.

Table 1

Different categories of the participants

Group	1	2	3	4	5
Age	<15	15-30	30-45	45-60	60-75

In our survey, we consider the following hypothesis,

H_4 Customer age positively influences customer churn.

2.5. Customer status

In this survey, customer status is considered as mediation effect and we have considered a customer active when the customer had, at least, one call within the last two months.

H_5 Customer status positively influences customer churn.

Based on different hypotheses accomplished for the proposed study of this paper, we have consider the following model,

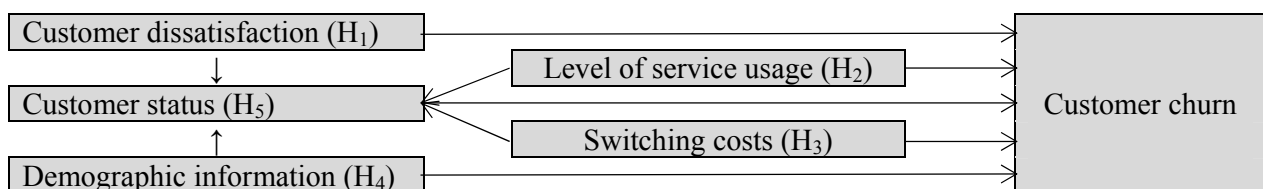


Fig. 2. The framework of the proposed study

2.6. Predicted data

The proposed study of this paper considers historical data from a group of 3150 customers who used Mobile phone in Iran. When a customer does not make a phone call for a period of two-month or sell his/her SIM card to another person, customer churn occurs.

2.7. Censored data

In some experiments, it is not possible to complete the test and in our case when we face with such incident and customer churn does not happen we say right-censored data happens.

2.8. Preprocessing

In this survey, we use different hidden layers and epoch to remove the existing noise in the data. Our experiments indicate that we could get the best results for two groups of churn and non-churn with 16 hidden layers and 100 repeats. Table 2 shows details of our findings in this stage.

Table 2
The results of preprocessing the data

Layer Hidden	Epoch			
	50	100	200	300
8	91.68	90.98	92.55	91.17
12	89.56	92.33	90.01	89.79
16	88.98	94.23	88.08	90.52
20	90.49	91.61	87.57	89.39
24	92.74	90.88	88.09	88.41
28	91.44	87.69	89.23	89.57
32	86.23	86.21	86.19	86.47

3. The results

In order to extract the customer churn probability we have used Cox regression analysis using SPSS software package and Table 3 demonstrates the results of our findings,

Table 3
The results of Cox regression analysis

Factors in the equation	B	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
					Lower	Upper
Call Failure	0.121	1	0.00	1.13	1.11	1.15
Complains	2.367	1	0.00	10.66	9.44	11.88
Charge Amount	-0.271	1	0.001	0.736	0.663	0.863
Seconds of Use	0.001	1	0.002	1.001	1.000	1.002
Frequency of use	-0.075	1	0.00	0.93	0.90	0.96
Frequency of SMS	-0.022	1	0.00	0.98	0.96	1.00
Distinct Called Numbers	-0.014	1	0.52	0.99	0.97	1.01
Age Group		1	0.00			
Age Group (1)	-12.41	1	0.83	0.00	0.00	4.06E-5
Age Group (2)	3.61	1	0.00	37.44	10.02	98.87
Age Group (3)	2.02	1	0.00	7.54	3.31	63.77
Age Group (4)	2.98	1	0.00	19.69	6.14	62.04
Internet subscription	-1.37	1	0.009	0.25	0.13	0.37
Status	0.42	1	0.004	1.52	1.14	2.02

As we can observe from the results of Table 3, except two cases, in all other cases, we can accept the impact of factors on customer churn. Table 4 and Table 5 demonstrate the results of testing various hypotheses.

Table 4
The results of testing different hypotheses of the survey

Independent variable	Hypothesis	Result
Call Failure	H _{1a}	Accept
Complains	H _{1b}	Accept
Charge Amount	H _{2a}	Accept
Seconds of Use	H _{2b}	Reject
Distinct Called Numbers	H _{2c}	Reject
Frequency of SMS	H _{2d}	Accept
Frequency of Use	H _{2e}	Accept
Internet subscription	H ₃	Accept
Age group	H ₄	Accept
Status	H ₅	Accept

Table 5
The results of intermediate hypotheses

Variable	Hypothesis	Result	Mediation effects
Call Failure	H _{1a'}	Reject	No mediation
Complains	H _{1b'}	Reject	No mediation
Charge Amount	H _{2a'}	Reject	No mediation
Seconds of Use	H _{2b'}	Reject	No mediation
Distinct Called Numbers	H _{2c'}	Reject	No mediation
Frequency of SMS	H _{2d'}	Accept	Partial mediation
Frequency of Use	H _{2e'}	Reject	No mediation
Tariff Plan	H _{3'}	Reject	No mediation
Age group	H _{4'}	Reject	No mediation

In order to compare the performance of the proposed model, we have used five criteria including prediction accuracy, errors' type I and II, root mean square error and mean absolute deviation and Table 6 shows the hypotheses for testing the effectiveness of the survey,

Table 6
The hypotheses of the survey for testing the proposed hybrid method

Predicted		Actual	
		Non-churners	Churners
Non-churners	Non-churners	<i>a</i>	<i>b</i> (II: $S^* > \alpha$)
	Churners	<i>c</i> (I: $H^{**} > \alpha$)	<i>d</i>

*Probability of Survival
**Probability of Hazard

where RMSE and MAD are calculated using the following,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_{ch}} (S_{ch}^i - 0)^2}{N_{ch}} + \frac{\sum_{j=1}^{N_{Nch}} (1 - S_{Nch}^j)^2}{N_{Nch}}}, \tag{3}$$

$$MAD = \frac{1}{N_{ch}} \sum_{i=1}^{N_{ch}} |e_{ch}^i - \bar{e}_{ch}| + \frac{1}{N_{Nch}} \sum_{j=1}^{N_{Nch}} |e_{Nch}^j - \bar{e}_{Nch}|, \tag{4}$$

where S_{ch}^i and S_{Nch}^j are survival value for churn and non-churn customers, respectively. In addition, $e_{ch}^i = (S_{ch}^i - 0)$ and $e_{Nch}^j = (1 - S_{Nch}^j)$, and finally, $\bar{e}_{ch}, \bar{e}_{Nch}$ are calculated based on the average values of e_{ch}^i and e_{Nch}^j , respectively. Table 7 shows details of our findings on the results of our survey,

Table 7
The results of comparing the performance of Pure Cox and Hybrid method

	Performance Metrics				
	Accuracy	Error type I	Error type II	RMSE	MAD
Pure Cox	82%	93	52	0.132	0.106
Hybrid method	96.66%	27	11	0.041	0.073

As we can observe from the results of Table 7, the hybrid method performs better than pure Cox method.

Finally, Fig. 3 and Fig 4 show details of our testing the effect of age on customer churn and survival, respectively. As we can observe from the results of Fig 3 and Fig 4, the highest customer churn belongs to people aged 15-30, which signals for some appropriate plans and strategies for customer retention on this group.

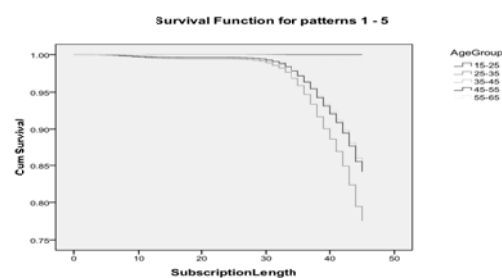
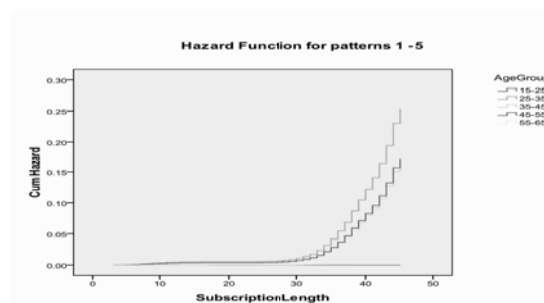


Fig. 3. Customer churn for different age groups **Fig. 4.** Customer survival for different age groups

4. Conclusion

In this paper, we have presented a hybrid method to measure the effects of various factors on customer churn. The proposed model of this paper first used neural network to learn the model and then implemented Cox regression function to predict customer behavior. We have compared the performance of the hybrid method versus pure Cox regression function and the results indicate that the proposed model of this paper outperforms the pure method based on various criteria.

Acknowledgment

The authors would like to thank the anonymous referees for constructive comments on earlier version of this work, which have significantly improved the quality of the work.

References

- Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., & Dedene, G. (2002). Bayesian neural network learning for repeat purchase modeling in direct marketing. *European Journal of Operational Research*, 138(1), 191–211.
- Berry, M. J. A., & Linoff, G. S. (2003). *Data mining techniques: For marketing, sales, and customer support*. John Wiley & Sons.
- Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications*, 34(1), 313–327.
- Gerpott, T. J., Rams, W., & Schindler, A. (2001). Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. *Telecommunications Policy*, 25, 249–269.
- Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecomm churn management. *Expert Systems with Applications*, 31(3), 515–524.
- Hung, C., & Tsai, C., F. (2008). Segmentation based on hierarchical self-organizing map for markets of multimedia on demand. *Expert Systems with Applications*, 34(1), 780–787.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis estimating mediation in treatment evaluations. *Evaluation review*, 5(5), 602–619.
- Keramati, A., Ardabili, S., M., S. (2012). Churn analysis for an Iranian mobile operator. *Telecommunications Policy*, 35(4), 344–356
- 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.
- Kim, H.-S., & Kwon, N. (2003). The advantage of network size in acquiring new subscribers. *Information Economics and Policy*, 15(1), 17–33.
- Madden, G., Savage, S.J., & Coble-Neal, G. (1999). Subscribers churn in the Australian ISP market. *Information Economics and Policy*, 11(2), 195–207.
- Ngai, E. W. T, Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36(2), 2592–2602.

- Pendharkar, P., C. (2009). Genetic algorithm based neural network approaches for predicting customer churn in cellular wireless network services. *Expert Systems with Applications*, 36, 6714–6720.
- Seo, D., Ranganathan, C., & Babad, Y. (2008). Two-level model of customer retention in the US mobile telecommunication service market. *Telecommunications Policy*, 32(3–4), 182–196.
- Tsai, Ch., & Lu, Y. (2009). Customer churn prediction by hybrid neural networks. *Expert Systems with Applications*, 36, 12547–12553.
- 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.