

# Model to Predict the Behavior of Customers Churn at the Industry

Keyvan Vahidy Rodpysb

Department of e-commerce, Nooretouba University, Tehran, Iran

## ABSTRACT

In order to continue life-sustaining competitive advantage, many organizations focus on maximizing the marketing relationship with their customer lifetime value and customer churn management. In fact, more organizations are realizing that their most valuable resource is their current customer base. In the present study are to go through a database collected from 300 customers, including an insurance company in Iran has been used. In order to check the model presented with a desire to review a decision tree classification methods (C5.0, CART, CHAID, and Quest), Bayesian networks and neural networks will be paid with respect to sample. Survey results can help managers, marketers in this arena is in various industries. Reduction strategies appropriate to offer in this field. The entire paper must be in A4 size and “Moderate” margin

## Keywords

Data Mining, Classification method, Decision tree, Customer churn, Insurance.

## 1. INTRODUCTION

Compression of marketing campaigns with marketing companies to get business from a major bus and the most valuable thing for companies is that customer paid. In this regard, in most industries, companies in order to fully realize the customer-oriented business strategies for maintain its competitive advantage creating and sustaining the level of their profits and revenue; retain customers based on their current. Offers a variety of advertising methods and provide special facilities to customers, especially the industry in terms of workload have more extent than be more found. [5] Customer churn is one of the threads in this field. Behavioral patterns of customers churn by the existing data is something that is long lasting in some industries such as telecommunications, banking, journalism, film industry, retail industry has taken place[2].

Thus, in order to predict churn customers can suggest some strategies help to Companies in every industry to identify any scope to turn churn customers through the adoption of appropriate strategies in order to take steps to deal with this phenomenon. One of the most valuable tools for knowledge extraction from data is data mining tools. Data mining is a process that uses intelligent techniques to extract knowledge from data set. Knowledge extracted in the form of models, patterns or rules provided these patterns, models and rules to provide various forms of knowledge can be extracted [2]

The fields used in this field can be mentioned the following:

- Using existing data mining models to predict customer churn [4]
- Evaluation of data mining models to predict customer churn [3]
- Model for predicting customer churn [2]

In this article, users are optimized to provide a model for predicting customer churn paid, in this context to examine the optimal model to evaluate this model with decision tree

classification methods (C5.0, CART, CHAID, and Quest), Bayesian networks and neural networks.

## 2. LITERATURE REVIEW

Churn important issue, because the lack of customers should be provided through new customers that it would cost too much for us. This issue, especially in industries that market is saturated, there will be more. Everyone wants to achieve through its product or service to attract new customers from the transition of customers to competitors. Everyone wants to achieve through its product or service to attract new customers from the transition of customers to competitors. Summary attract new customer costs 5 to 6 times the cost of maintaining existing customers [2]

Administrators can use the tools to better analyze issues in order to determine customer churn reduction strategies to help deal with it is data mining. Data mining is a tool to discover patterns and trends with significant dependencies and checking a large amount of data stored in the storage using pattern recognition techniques with mathematical and statistical methods. [1]

As seen in Table 1, modeling techniques to predict the frequency churn customers using data mining these include methods such as decision trees, logistic regression, neural networks, Bayesian networks, random forests, association rule, support vector machines modeling capability to provide customer churn with data analysis is created. Research literature in order to understand the reasons for the low hand churn customer [15]

Today, most businesses have realized the importance of data mining because it is well know that through it we can identify the characteristics and behavior of their customers and companies with their need to regulate and prevent their transition to the company.

## 3. TYPESET TEXT REASEARCH METHODOLOG

Figure 1 presents the structure for the project is shown in the first step to examine models of decision trees (C5.0, CART, CHAID, Quest), Bayesian networks and neural networks will be tested on sample data. Then the second step, the model presented by this study on the sample data we tested.

### 3.1 Case study

Experimental data used in this study from insurance companies in time interval of 23JUL-23SEP 2011. Insurance firms in the city of Anzali, Gilan province in Iran has been prepared. This data includes 300 samples from customers in the insurance branch of the Iranian city of Bandar Anzali In this study, eight variables (answer questions / problems, Customer satisfaction of the insurer's behavior, The Company's obligations to pay compensation, Iran Insurance Company of premiums, How to pay premiums to the insurer, Credit insurer or insurance company in Iran, Find out which insurance companies use insurance and Career Cluster). Churn the index as an indicator of this variable to variable will be dealt with.

**Table 1. Application of data mining in customer churns the research literature**

Subject	Data mining techniques	Business	References
Analysis of customers churn	Association Rule Apriori	Banking	[4]chiang et al,2003
Identify and predict customer churn	Decision tree, support vector machines	Insurance	[16]Morik and Kopck,2004
Churn prediction model as part of a customer lifetime value model	Logistic regression, decision tree, neural network	Wireless telecom	[17]Hwang H., Jung 2004
Subject	Data mining techniques	Business	References
Evaluation of data mining methods to maintain current customers	Markov chain, Random forest, Logistic regression	Pay TV company	[3] Burez et al,2007
Model of Customer Churn Prediction on Support Vector Machine	neural network, decision tree, logistic regression, and Bayesian networks	Wireless telecom	[7]Guo-en and wei-dong,2008
An experimental study on four models of customer churn prediction	Bayesian networks, support vector machines	Wireless telecom	[8]Zhu et al ,2009
Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications	neural network, decision tree	Wireless telecom	[9]Huang et al, 2010
Building comprehensible customer churn prediction models with advanced rule induction techniques	C4.5, ant miner, support vector machines and logistic regression	Wireless telecom	[2]Verbeke et al,2011

### 3.2 Classification model used

In the present study was to evaluate the proposed model to evaluate the model will be addressed through existing classification methods. Hence, our sixth category decision tree method (C5.0, CART, CHAID, and Quest), Bayesian networks and neural networks are used for the charges.

#### Decision tree CART

CART decision tree and a recursive binary segmentation procedure is capable of processing particular attributes with continuous and discrete values. Trees have been grown without the use of any law to the greatest extent possible and then get to the root pruning. Divide that into the next case will be pruning trees in overall performance on the test data plays a minimal role. More than one division at a time may be removed by this operation. The overall goal of this algorithm

is a series of pruning trees and representation but each of them is candidates for optimal trees. Tree of appropriate size by calculating the predicted performance of each tree in the pruning process is determined. [1]

#### Decision tree QUEST

QUEST stands for Quick, Unbiased, and Efficient Statistical Tree. It is a relatively new binary tree growing algorithm .It deals with split field selection and split-point selection separately. The univariate split in QUEST performs approximately unbiased field selection. That is, if all prediction fields are equally informative with respect to the target field, QUEST selects any of the predictor fields with equal probability. QUEST affords many of the advantages of CART. You can apply automatic cost-complexity pruning to QUEST tree to cut down its size. QUEST uses surrogate splitting to handle missing values.

#### Decision tree CHAID

CHAID is a very effective statistical technique for segmentation, or tree growth. These methods of statistical values that have been evaluated as homogeneous values with consideration of factors combine and all other values of the heterogeneous (non-identical) are maintained. The recursive process continues until the tree is fully grown. CHAID is a binary tree in two ways. And can be more than two categories at any particular level of the tree will produce. Because the trees tend to grow larger than binary methods.[1]

#### Decision tree C5.0

Decision tree algorithm C5.0 algorithm new features such as variable costs of providing classified. For classification C5.0 algorithm to minimize the expected cost of classification error rates, rather than making. There is even the possibility that cases are not presented as of equal importance.C5.0 algorithm provides facilities for the feature weight to calculate the amount each pays. With this feature, the algorithm attempts to reduce the error rate is predicted. In some recent applications of data mining, data volume is greatly increased. C5.0 Prior to classes, attributes can be automated screening and perform this function with omitted characteristics that are less dependent than other traits. [1]

#### Neural network

Neural networks are simplified models of neural fields human brain is designed for computers. So does the neural network learning that the records are tested then an estimate for each record is created every time an estimate was incorrect weights adjusted. I will continue this process to improve the estimation of conditional approval to be a powerful neural network estimators are estimating that the other methods as well and sometimes better than they do. [1]

#### Bayesian networks

Bayesian network based approach, based on mathematical rules through combining new information with existing knowledge and expertise is in accordance with the Bayesian network. General, Bayesian network graphical model to illustrate the possible relationships between a set of variables. When this model of expert knowledge, incomplete, ambiguous, and is associated with uncertainty [13] Bayesian network is defined by two factors:

- 1) Nodes that represent variables with a finite number of conditions which could indicate the type of each parameter measured, or the hypothesis of hidden variables.
- 2) The edges for a link between variables and conditional probability distribution of each variable to the variable condition of the marginal distribution of parents say Connection between the edges, they indicate the possible dependence, while the edges indicate that the causative relationship between two variables, so that the edges will not create any distance.

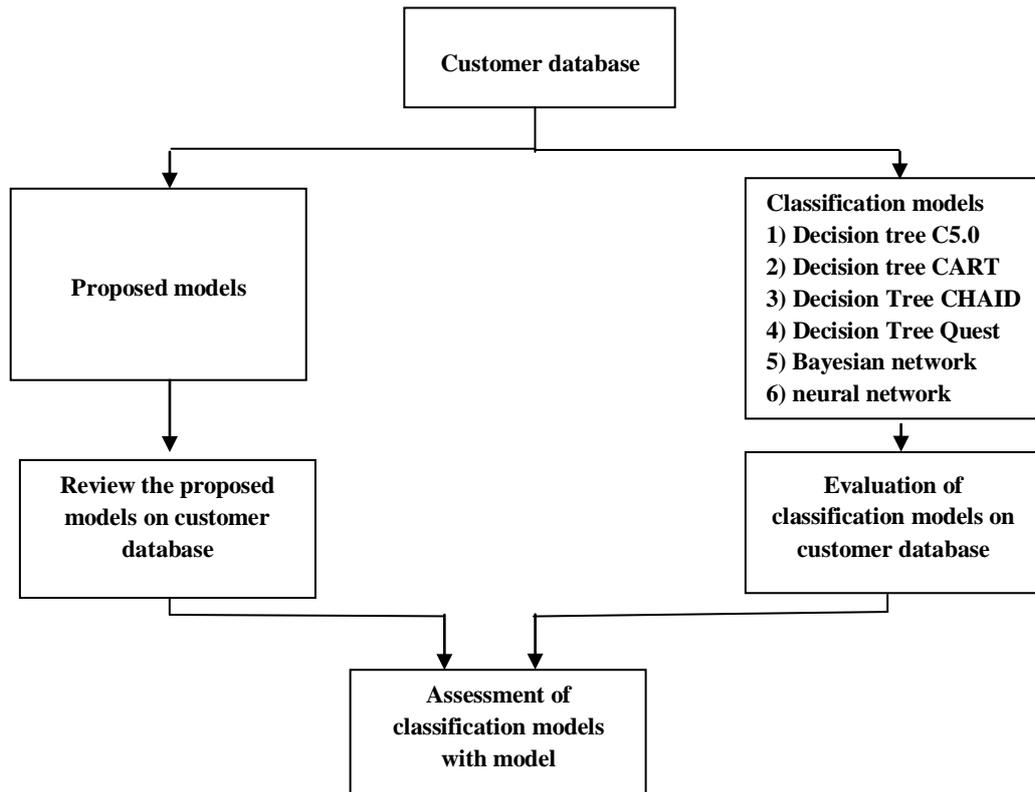


Fig. 1 .Research methodology

### 3.3 Proposed Model

Data mining model is presented for predicting customer churn model is a tree, we say that the normalization tree. Tree normalization process is as follows:

- 1) The input data is received
- 2) data compiled by the normalization of the MIN-MAX, in the interval 0 to 1 is normal.
- 3) Each customer score variables with n variables are normalized to the interval 0 to 1 as  $m_1$  to  $m_n$  is calculated as follows.

$$S_c = \frac{1}{n} \sum_{i=1}^n m_1 + m_2 + \dots + m_n$$

- 4) Next, in order to determine the index score for each variable to draw a tree with  $S_c =$  Customer score,  $X_c =$  Customer behavior and  $N =$  number of customer, we perform the following

$$Z_m = \frac{1}{N} \sum_{i=1}^N X_1 S_1 + X_2 S_2 \dots + X_N S_N$$

After the decision tree model on the normalization of data is determined according to objective criteria. The effect of these variables is between 0 and 1

### 4. IMPLEMENTATION

I mentioned before in this study. Action to review a decision tree classification method (C5.0, CART, CHAID, and Quest), Bayesian networks and neural networks to customers on the database most efficient method we apply. Table 2 Comparison of methods classified according to four indicators of the profit function, the overall accuracy, area under ROC curves and

Lift Index, please see. CART decision tree technique has better performance than other techniques. It seems unusual that this algorithm shows a better performance but due to data that are actually a set of results is not far-fetched. CART decision tree after the effect of data on indicators was evaluated in customers who are employees, the largest in this group of customers leaving that Although most activity is now Iran in the insurance, but dissatisfaction is important primarily to problems of accountability and they will encourage companies to transition. The customer engineer and doctor of

Table2. Binary classification evaluation techniques

Prediction method	Overall accuracy	Profit	ROC curve	Index Lift
decision trees CART	99.66	215	1	3.33
Bayesian networks	99.33	215	1	3.33
neural networks	99.66	210	1	3.33
decision trees Quest	99.33	205	1	3.33
decision trees C5.0	98.66	200	0.99	3.33
decision trees CHAID	98.66	200	0.99	3.33

The insurer is more important. The company's commitment for the customers of farmers and workers compensation insurer other than the degree of importance. And in the end market for customers that are most important for this group of customers the most effective agent for the insurer, the current level of commitment to the compensation, the amount of mutual trust, reputation and recognition. Now, after the tree normalization made on the data about the test results indicate that tree normalization relative to the decision tree CART

more attention to three factors of the company's commitment to open, compensation, satisfaction and customer behavior, the Insurer and amount of services insurance, the insurers in which one of the insurance companies of the importance .Higher. Evaluation model shows better performance than decision tree normalization relative to the decision tree CART is evaluated assuming the amount of expenses 5, earning 10 and weight 1 used

- Calculate the profit function

$$\text{Profit} = \sum_{m=1}^N F_j(t) * P_j(t)$$

$F_j(t)$  = Total amounts for all records in the frequency field t

$P_j(t)$  = Determined value for the catalog j

Fig 2. Shows a profit for the records as you can see the maximum amount is equal to 236.1607



Fig 2. Function to record profits

- The overall accuracy of the model

$$\text{Overall accuracy} = \frac{T_P + T_N}{N + P} = \frac{300 + 0}{300 + 0} = 1$$

- area under the curve: ROC curve for determining we proceed as follows:

$$x = TPrate(t) = \frac{300}{300} = 1$$

$$y = FPrate(t) = \frac{0}{300} = 0$$

$$A_{ROC} = \int_0^1 \frac{TP}{P} d \frac{FP}{N} = \int_0^1 1 = 1$$



Fig 3.exponential index of graphs Lift

## 5. CONCLUSION

As mentioned before, the model proved to be more efficient model, which examined data from the data collected from insurers in the Iranian cities has been discussed. And then to evaluate the decision tree classification methods (C5.0, CART, CHAID, and Quest) , Bayesian networks and neural networks. CART decision tree evaluation methods show good performance compared with other methods shows a binary classification. That seems to be verifiable. CART decision tree, because the data set (labeled) shows better performance. The results of the decision tree normalization in contrast to the results of decision tree CART decision tree above model and pay more attention to three factors of the company's commitment to open, compensation, satisfaction and customer behavior, the insurer and the amount you wish to use insurance services, in which the insurers of insurance companies is the degree of importance. The results show that the cases are considered experts in insurance and insurers less importance than the items that customers should have churn

## 6. REFERENCES

- [1] Berry ,Gordon S.linoff ,2004, , EBook Data Mining Technique for marketing Sales and CRM: Wiley Publishing, Inc., Indianapolis, Indiana
- [2] Verbeke , Martens , Mues , Baesens ,2011, Building comprehensible customer churn prediction models with advanced rule induction techniques, Expert Systems with Applications, Vol.38 , pp. 2354–2364
- [3] Burez, Van den Poel,2007, CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services”, Expert Systems with Applications, Vol.32 , pp.277–288H
- [4] Chiang, Wang , Lee , Lin,2003,Goal-oriented sequential pattern for network banking churn analysis”. Expert Systems with Applications, Vol 25 , pp.293–302
- [5] Tsai , Lu.,2009,Customer churn prediction by hybrid neural networks, Expert Systems with Applications, Vol.36,pp.12547-12553
- [6] Hwang, Jung , Suh , 2004,An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry,Expert Systems with Applications, Vol.26,2004, pp.181–188
- [7] Guo-en, Wei,2008, Model of Customer Churn Prediction on Support Vector Machine”, Systems Engineering Theory & Practice,Vol. 28,2008,pp. 71-77
- [8] Zhu, Qi, Wang .,2009,An experimental study on four models of customer churn prediction, Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on , Issue Date. 11-14 Oct 2009, pp.3199 – 3204
- [9] Huang , Buckley , Kechadi.,2010 b ,Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications ,Expert Systems with Applications Vol,37,pp.3638–3646
- [10] SPSS Inc,Clementine 12.0 Algorithms Guide , 2007
- [11] Jonathan Burez; Dirk Van den Poel. CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services, Expert Systems with Applications, (2007), vol.32, pp. 277-288

- [12] Burez, Van den Poel,2008,Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department, *Expert Systems with Applications*, Vol.35, ,2008, pp.497–514
- [13] Suebnukarn , haddawy,2006, A Bayesian approach to generating tutorial hints in a collaborative medical problem-based learning system , *Artificial intelligence in medicine*,Vol.38,pp.5-24
- [14] Coussement, Van den Poel, 2008., “Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques” , *Expert Systems with Applications* , Vol.34,pp.313–327
- [15] Chang Lee ,Yong Jo,2010,Bayesian Network Approach to Predict Mobile Churn Motivations: Emphasis on General Bayesian Network, Markov Blanket, and What-If Simulation, *Future Generation Information Technology Lecture Notes in Computer Science*,Vol.6485,2010,pp. 304-313
- [16] Morik , opcke ,2004, Analyzing Customer Churn in Insurance Data, *Knowledge Discovery in Databases: PKDD 2004 Lecture Notes in Computer Science*Vol.3202, pp.325-336
- [17] Hwang, Jung , Suh , 2004 , An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, *Expert Systems with Applications*, Vol.26,2004, pp.181–188