2015 International Conference on Computational Intelligence and Communication Networks

Application of Computational Intelligence to predict churn and non-churn of customers in Indian Telecommunication

Ramakanta Mohanty
Keshav Memorial Institute of Technology, Narayanaguda, Hyderabad, India
ramakanta5a@gmail.com

Jhansi Rani K
Keshav Memorial Institute of Technology, Narayanaguda, Hyderabad, India
jhansi.hema25@gmail.com

***Abstract*—In the modern society, mobile communication became the leading medium of communication. Now the public policies and standardization of mobile communication allows customers to switch from one service provider to another service provider easily. One of the most critical challenges in data and voice telecommunication service industry is retaining customers. The cost of retaining an existing customer is lesser than the cost of getting a new customer. So service providers now shifted their focus from customer acquisition to customer retention. As a result, churn prediction has emerged as the most essential Business Intelligence (BI) application that aims to identify the customers who are about to transfer their service to a competitor i.e. to churn. In this paper, we proposed Counter Propagation Neural Networks (CPNN), Classification and Regression Trees (CART), J48 and fuzzyARTMAP to predict customer churn and non-churn in telecommunication sector. The dataset analyzed is taken from Indian Telecommunication Service Industry.**

Keywords: fuzzyARTMAP, Counter propagation Neural Network (CPNN), Churning and Non-churning, Classification and Regression Trees (CART), J48.

I INTRODUCTION

Banks, telephone service companies and internet service providers, etc. often use customer churn analysis and rates which play a vital role in business metrics [1-5]. Over the two decades, the customers of telecom service provider are increasing day by day. The service provider provides different services to customers through Mobiles, Landlines, Wireless Phones and Internet Leased Line. In the telecom industry, existing customers can choose any multiple service providers and also they can easily switch from one service provider to another service provider. In this competitive industry, customers can demand products and services at lower prices while the service providers focus on the development of their business goals. The factors that influence occurring of customer churn are due to variation in prices, Lack of customer services, billing errors, lack of connection capability, lack of network coverage, entering of new service providers and lack of technology [6-15]. To overcome customer churn every industry should concentrate on the factors that influences customer churn. This paper describes about predicting of customer churn in telecommunication dataset.

Customer churn is also known as customer attrition, customer defection, or customer turnover [5, 16, 18, 19, 20]. Churn is defined as the loss of customers. There are two types of churns. One is voluntary churn and the other is involuntary churn. The decision taken by the customer to switch from one company to another company or service provider is called voluntary churn. Customer relocation for long term, death and transfer of job are involuntary churns. Generally, analysts concentrates on voluntary churn which affects the relationship between company and customer such as better products, better delivery methods, lower prices, building satisfactory customer relationships, better marketing and, above all, successful customer communications.

In this paper, we propose Counter Propagation Neural Network, CART, J48 and fuzzyARTMAP to predict customer churn in the telecommunication sector and we collected data from Indian telecommunication sectors.

The rest of the paper is organized in the following manner. A brief discussion about Literature Survey is presented in section 2. Section 3 presents the overview of counter propagation neural network, fuzzyARTMAP, CART and J48 techniques. Methodology of the paper is presented in section 4. Section 5 presents a detailed discussion of the results and discussions. Finally, section 6 concludes the paper.

II LITERATURE SURVEY

In recent years, Customer churn becomes a focal concern in Indian telecommunication markets due to the saturated markets and competitive business environment. The customer churn is not limited to just telecommunication, banking and financial industries but also prevalent in other service industries such as mobile telecommunication, television viewership, etc. We present a brief overview of the literature involving customer churn prediction in different domains as well as the applications of fuzzyARTMAP and counter propagation neural network (CPNN) in diverse areas of science and technology. Some of the works reviewed discuss various models developed for the problem, whereas some other discusses the managerial implications of churn problem. Au et al. [1] employed data mining by evolutionary learning on Credit Card Database and their experimental results showed that the revolutionary learning is a robust way to predict the churn.

Hu [2] developed an ensemble model consisting of Bayesian network, neural network and decision tree on third party data, segmentation files and payment database. He found that

978-1-5090-0076-0/15 $31.00 © 2015 IEEE
DOI 10.1109/CICN.2015.123

ensemble model predicts better accuracy than a standalone method for churn prediction.

Larivière and Poel [3] used European financial services firm dataset by employing random forest and linear regression. From their simulated study, they found that Random forest provided better accuracy than linear regression. Hadden et al. [4] simulated the neural networks combined with a GA-based rule discovery method to predict the customer churn with high accuracy.

Abbasimehr et al. [5] employed C 4.5, Ripper and PART to predict the churning of customer. In their simulated results Ripper outperforms in terms of accuracy and comprehensibility. Hossein et al. [6] employed ANFIS-Subtractive (subtractive clustering based fuzzy inference system (FIS)) and ANFIS-FCM (fuzzy C-means (FCM) based FIS) models to analyze the performance in terms of accuracy, specificity and sensitivity of churn prediction applications. Kaur et al. [7] developed the churn generators for testing their overlay application in the oversim simulation environment. Zhang et al. [8] proposed a behavior-based telecom customer churn prediction system which uses only customer service usage information to predict customer churn using a SOM clustering algorithm. Forhad et al. [9] employed unsupervised learning technique called Rule-based learning to predict whether a customer will churn in the near future or not using billing data of a telecom company. Qinet al. [10] introduced OLAP analytic process of some telecom churn customer. They have proven that because of the multidimensional characteristic of OLAP, it can provide users with a multi-perspective and multi-level description model. Zhang et al. [11] employed the prediction models that are built using machine learning algorithms. They use a customer data set of a Chinese mobile telecommunication. Three machine-learning methods are adopted including logistic regression (LR), decision tree (DT) and neural network (NN). The tool used for training the models was SAS Enterprise Miner and they measured the effects of network attributes on prediction accuracy. Pushpa et al. [12] proposed Telecom Social Network Analysis (TSNA) for identifying structural and behavioral patterns in systems based on the relations between customers. They have mined churners and non-churners communities by using a Regular equivalence algorithm for multi-relational telecom network. Jinboet al. [13] used the AdaBoost to predict the customer churn on credit debt customer database from commercial bank in China. They used three different boosting schemes i.e.Real AdaBoost, Gentle AdaBoost and Modest AditaBoost and compared the results obtained by them with SVM. Qiet al. [14] proposed the TreeLogit model, which integrates the advantage of AD tree model and logistic regression model to improve the predictive accuracy and to interpret the ability of the churn prediction model. Salman et al. [15] employed multi-dimensional time based analytical approach to capture the churn happening any time in the life cycle of a subscriber and they employed Logistic Regression as a primary classification technique.

Hung et al. [16] employed decision tree and multilayer perceptron on a wireless telecom company customer data which included customer demographics, billing information, contract/service status and call detail records and service change log. Mutanen [17] used logistic regression on personal retail banking to analyze the churning of customer. Neslin et al. [18] applied logistic regression and decision trees, anonymous wireless telephone carrier data and they found that Logistic regression and decision tree provided better accuracy. Chu et al. [19] developed a churn model in telecommunications data using decision tree and achieved an accuracy of 85%. Wezel and Potharst [20] worked with real-life marketing datasets and developed an ensemble model consisting of decision trees and logistic regression and results obtained by them outperformed the decision tree and logistic regression. Kumar and Ravi [21] conducted an investigation on the credit card churn prediction problem in bank credit cards by presenting ensemble system with majority voting multilayer perceptron, logistic regression, decision tree (J48), random forest, radial basis function network and support vector machine and result obtained by them with a sensitivity of 92.37%. Most recently, Naveen et al. [22] employed fuzzyARTMAP to predict the customer in banking sectors by using a Latin-American bank dataset.

Therefore, it is clear that fuzzyARTMAP and CPNN has not yet been applied to solve problems in telecommunications. This paper precisely addresses that gap. Secondly, application of fuzzyARTMAP marks the beginning of the application of a tightly coupled soft computing system to customer churn prediction problems concerning telecommunication sectors. Finally, the impact of feature selection on fuzzyARTMAP and CPNN in churn prediction is an area of active investigation. These points form the motivation for the present study.

III OVERVIEW OF TECHNIQUES EMPLOYED

A. Counter Propagation Neural Network (CPNN)

The Counter propagation Neural Networks (CPNN) is a multilayer network consists of input, clustering and output layers which was introduced by Hecht-Nielsen [23]. It is a combination of self-organizing map of Kohonen and Grossberg layer. It is trained in two stages. In stage 1, the input vectors are clustered based on Euclidean distances or dot product method and in stage 2, the desired response is find out by adopting the weights from the cluster units to the output units. CPNN functions in two modes i.e. normal mode and training mode. In normal mode, the input vector is passed to input of CPNN and output vector is obtained. But, in case of training mode, the input vector is given to the input of the CPNN and the weight are adjusted to obtain desired output vector.

B. fuzzyARTMAP

The fuzzy ARTMAP is an unsupervised neural network that performs supervised learning for recognising categories and multidimensional maps in response to input vectors (Carpenter et al., [24]. Earlier adaptive resonance theory models consisted of ART1 and ART2 of learning systems which classify inputs by fuzzy set features with membership value between 0 and 1. This learning system is achieved by replacing the ART1 modules (Carpenter and Grossberg, [24, 25] of the binary ARTMAP system with fuzzy A$\bar{R}$T modules (Carpenter et al., [22], [25]).The fuzzyART neural network consists of two layers of nodes called F_1 and F_2. F_1 represents an input layer and F_2represents the competitive layer. A set of real valued weight $W = \{w_{ij} \in [1,0]: i = 1,2.....,M\}; j = 1,2...,N\}$ is related to F_1 to F_2 layers. In F_2 node, j represents a recognition category that learns a prototype vector $w_j= (w_{1j}, w_{2j}...,w_{mj})$.*F2* layer is connected through learning associative links to an L node map field F^{ab}, where L is the number of classes in the output space. A set of binary weights $W^{ab}= \{w_{jk}^{ab} \in \{0,1\}; j = 1,2...N; k = 1,2,...,L\}$ is associate with the F_2 to F^{ab} connections. The vector $w_j^{ab} = (w_{j1}^{ab}.w_{j2}^{ab}....w_{jL}^{ab})$ links F_2 node j to one of the L output classes. During training, ARTMAP performs mapping between training set vectors $a=(a_1, a_2,...,a_m)$ and output levels $t=(t_1, t_2, ...,t_L)$, *where* $t_k=1$ *if* K is the largest class label for a and zero elsewhere(Granger et al., (21)). The architecture of fuzzyARTMAP is depicted in Fig. 1.

C. J48

The J48 decision tree is an algorithm based on C 4.5 that was developed by J. Boss Quinlan [27]. In order to classify, a decision tree is created based on the attribute values of training data. When the training set is encountered, an attribute that discriminates the various instances is easily identified. This feature is called information gain and it is used to classify the data. Among the different possible values of this feature, if there is any value for which there is no ambiguity, the branch is terminated. Then, the target value obtained by that branch is associate to it. Further, any other attribute that provides the highest information gain is searched. The iterations continues till the decision of combination of attributes provide a target value is obtained or all attributes exhausted. The J48 is available in Knime Software tool [26].

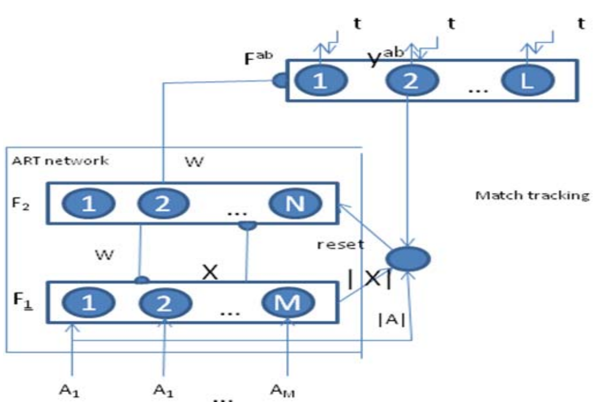

Fig.1 Architecture of fuzzyARTMAP for pattern classification [24]

D. CART

Classification and regression tree [28] is a classification/regression model which is used to construct a decision tree by using historical data. CART was developed by Breiman et al. [29]. Decision trees are represented by set of question which split the training data into smaller and smaller parts, resulting in a set of "If-then rules". These rules can be used to solve classification and regression problems. Decision tree consists of three parts i.e. construction of tree, choice of right tree, and classification of data by using constructed tree. Ginni and Twoing splitting rule is used to construct the tree. To choose the right size of the tree, tree optimization method is used by cutting off the insignificant nodes and subtrees. The import factors of CART analysis are the set of rules for splitting of each node into a tree, deciding which tree is incomplete and assign each leaf node to a class outcomes or prediction value for a regression.

IV. METHODOLOGY

In this paper, we collected the data from Indian telecommunication sectors. The datasets consists of 125 samples and five attributes. The dataset is divided into three folds. We performed a 3 fold cross validation in the entire experimentation. The different attributes of the dataset are customer dissatisfaction, switching costs, service usage, customer related variables, and customer status.

(i) Customer dissatisfaction:

In Indian telecommunications, the key factors for customer satisfaction or dissatisfaction are network quality and call quality. The factors that influence customer churn are call drop rates, call failure rates and number of complaints.

(ii) Switching Costs

Switching costs are related to constraints or loyalty. It depends upon the service providers to satisfy the customers in terms of quality of service and service contact. The switching costs also related to accumulated loyalty points and loyalty program membership.

(iii) Service usage:

This is basically how a customer uses the service provided by service provider in terms of used measures, minutes of use, frequency of use and total number of receivers contacted by the subscriber. The customer tends to churn if any one of uses does not meet properly. Further, it depends on monthly billed amount, unpaid balances and unpaid monthly bills.

(iv) Customer related variables:

Customer related variables are customer grade, calling plans, gender, payment method, handset capability and hand set manufacturer.

(v) Customer Status:

The customer status can be active use, nonuse and suspended. The customers are generally churn when the status is nonuse or suspended.

In this paper, we are dealing with binary classification problems. The term sensitivity, specificity and accuracy used to classify the customer into churn and those who do not churn. The quantities employed to measure the quality of the classifiers are sensitivity, specificity and accuracy, which are defined as follows [30]:

Sensitivity is the proportion of the churn positives, which are correctly identified.

Sensitivity = $\frac{True\ Positive}{(True\ Positive + False\ Positive)}$

Specificity is the proportion of churn negatives, which are correctly identified.

Specificity = $\frac{True\ Negative}{(True\ Negative + False\ Positive)}$

Accuracy is the proportion of true results, either churn positives and churn negatives, which are correctly identified.

Accuracy=

$$\frac{True\ Positive + True\ Negative}{(True\ Positive + True\ Negative + False\ Positive + False\ Negative)}$$

Figure 2 shows the confusion matrix of customer churn. Several standard terms have been defined the two class matrix. There are four terms such as True Positive (TP), True Negative (TN), False Negative (FN) and False Positive (FP). If a customer churns, the predicted result is true positive. Similarly, if the customers do not churn, the predicted results are true negative. Both true positive and true negative suggests consistent results between predicted and actual conditions. If the predicted results indicate the customer to churn but customer is loyal, the result is false positive. If the experimental result predicted that the customers do not churn, the result is false negative.

V. RESULTS AND DISCUSSIONS

We employed CPNN, FuzzyARTMAP, CART and J48 techniques for our experiments. We followed three fold cross validation and each fold is given as input to Counter Propagation Neural Network (CPNN), CART, J4.8 and fuzzyARTMAP. We employed "Counter-propagation neural networks" in Matlab environment to carry our experiments. Two third (2/3rd) of samples of dataset have been used for training the network and one third (1/3rd) of the samples of dataset have been used for testing. From our experiment, we observed that, the average sensitivity of 94.03% obtained in CART and followed by FuzzyARTMAP of 91.07%.The average sensitivity of CPNN and J48 is 89.03% and 82.07%, respectively, but in fold2 of CPNN obtained sensitivity of 100% which is a remarkable result in our study presented in Table 1.

		Predicted: True	Predicted: False
Actual	True	True Positive (TP)	False Negative (FN)
	False	False Positive (FP)	True Negative (TN)

Fig.2. Confusion Matrix

Table 1 Sensitivity (%) in different folds and average Sensitivity(%)

Techniques	Fold1	Fold2	Fold3	Average
CPNN	92.30	69.23	69.23	76.92
fuzzyARTMAP	100	100	95	98.33
CART	77.05	77.05	77.05	77.05
J4.8	37.5	81.25	93.75	70.83

From our simulation study, we observed that average sensitivity in the case of CPNN is 89.83%. If we compare three different fields, fold2 resulted 100% sensitivity compared to other two folds. The average sensitivity of CART is 94.03% and an average sensitivity of J4.8 is 82.07%. The fuzzyARTMAP is also resulted average sensitivity of 92.08%. Out of four techniques, CART outperformed all other techniques in terms of sensitivity is concerned.

From Table 2, we observed that the average specificity is 98.33% in fuzzyARTMAP followed by CART of 77.05%. The average specificity of CPNN and J48 is 76.92% and 70.83, respectively. In fuzzyARTMAP, fold1 and fold2 obtained 100% specificity.

In Table 3, the accuracy of fuzzyARTMAP resulted 97.11%, which is better than any other techniques and the other folds. As long as accuracy is concerned, CPNN and CART resulted approximately same results and J48 resulted the least results. We obtained highest accuracy of 100% in fold2 in case of fuzzyARTMAP. Here, fuzzyARTMAP is outperforming all other techniques especially in fold2 which is a significant result as long as accuracy is concerned.

We developed the code for fuzzyARTMAP in 'C' language and perform the simulation. In fuzzyARTMAP 70% of the samples have been used for training and 30% of the samples have been used for testing. The parameters which are used in fuzzyARTMAP are shown in Table 4. With these parameters, we carried our experiments for all the folds. The results for the three folds are listed in Table 2.

Table 2 Specificity (%) in different folds and average Specificity(%)

Techniques	Fold1	Fold2	Fold3	Average
CPNN	83.33	88.88	83.33	85.18
fuzzyARTMAP	97.22	100	94.11	97.11
CART	86.95	86.95	86.95	86.95
J4.8	65.78	78.94	87.17	77.29

Table 3 Accuracy (%) in different folds and average Specificity (%)

Techniques	Fold1	Fold2	Fold3	Average
CPNN	78.20	100	91.30	89.83
fuzzyARTMAP	93.75	93.75	87.5	91.67
CART	94.03	94.03	94.03	94.03
J48	86.36	77.27	82.60	82.07

Table 4 Parameters used in fuzzyARTMAP

Dimensionality of the input space	5
Dimensionality of the output space	2
Number of training patterns	81
Number of test patterns	34
Total number of patterns	115
Alpha (α)	0.001
Epsilon (ε)	0.001
ART-A rho (ρ)	0.0
ART-B rho (ρ)	1.0

We conducted further investigations on the dataset in order to come out with a 'rule-based' expert system. This is where the rules extracted by CART and J48, presented in Tables 5 and 6 respectively. A close look at the rules induced by both CART indicates that customer related variability dominates the entire scene and the rules do not depend on any other attribute. In case of J48, customer related variables, switching costs and customer status are dependent on customer's churning or non-churning behavior. From the Table 5 and 6, we observed that customer related variables play a vital role for churning of customer. Though, under customer dissatisfaction i.e. call drop rate, call failure rate and the number of complaints plays an equal role which leads to customer related variations.

Table 5 Rules extracted by CART for the three folds cross validation

Rule No	Rule Antecedents	Classification
1	If x_4<=0.98, then	True
2	If x_4>0.98, then	false

Table 6 Rules extracted by J48 for the three folds cross validation

Rule No	Rule Antecedents	Classification
1	If x_4<=0.95 and x_2<=0.95 and x_2<=0.75 and x_5>0 and x_2<=0.725 then	True
2	If x_4<=0.95 and x_2<=0.95 and x_2<=0.75 then	True
3	If x_4<=0.95 and x_2<=0.95 and x_2<=0.75 and x_5>0 and x_2>0.725 and x_4<=0.5 then	True
4	If x_4<=0.95 and x_1<=0.304	False
5	If x_4<=0.95 and x_2<=0.95 and x_2<=0.75 and x_5<=0 and x_4>0.45 then	False
6	If x4<=0.95 and x2<=0.95 and x_2<=0.75 and x_5>0 and x_2>0.725 and x_4>0.5 then	False
7	If x_4<=0.95 and x_2>0.95 and x_5<=0.429 then	False

x_1-Customer Dissatisfaction, x_2-Switching Cost, x_3-Service Uses, x_4 Customer related variables, x_5-Customer Status

VI. CONCLUSIONS

In this paper, we employed CPNN, fuzzyARTMAP, CART and J4.8 to predict the churning and non-churning of customers in the Indian telecommunication market. From our simulated results, CART outperformed all the techniques in terms of sensitivity and followed by fuzzyARTMAP. FuzzyARTMAP provided best specificity and accuracy compared to other techniques and followed by CART. In practice, to predict churning and non-churning in telecommunication or in the banking sector, we suggest to use fuzzyARTMAP and CART rather than other techniques as it provides spectacular results. This is significant outcome of this study.

REFERENCES

[1] R. N. Bolton, P. K. Kannan, and M. D. Bramlett, "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," Journal of the Academy of Marketing Science, vol. 28, No. 1, pp. 95–108, 2000.

[2] W. H. Au, K. C. C Chan, and XinYao, "A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction," IEEE Transactions on Evolutionary Computation, vol. 7, No. 6, pp. 532–545, 2003.

[3] X. Hu , "A Data Mining Approach for Retailing Bank Customer Attrition Analysis," Applied Intelligence, vol. 22, No. 1, pp. 47–60, 2005.

[4] B. Larivière, and D.Van den Poel, "Predicting Customer Retention and Profitability by using Random Forests and Regression Forests Techniques," Expert Systems with Applications, vol. 29, No. 2, pp. 472–484, 2005.

[5] J. Hadden, A. Tiwari, R. K. Roy, and D. Ruta, "Computer Assisted Customer Churn Management: State-Of-The-Art and Future Trends," Computers and Operations Research, vol. 34, No. 10, pp. 2902–2917, 2007.

[6] H. Abbasimehr, M. Setak, and M. J. Tarokh, "A Neuro Fuzzy classifier for customer churn prediction," International Journal of Computer Application, vol. 19, No. 8, pp. 35-41, 2011.

[7] R. Kaur, A.C. Sangal, and K. Kumar, "Analysis of Different Churn Models in Chord Based Overlay Networks," IEEE proceedings of RAECS, UIET, Punjab University, Chandigarh, India, pp. 1-6, 2014.

[8] Y. Zhang, R. Liang, Y. Li, Y. Zheng, and M. Berry, "Behavior-based Telecommunication Churn Prediction with Neural Network Approach," IEEE International Conference on Computer Science and Society (ICCS), Kota Kinabalu, pp. 307-310, 2011.

[9] N. Forad, M.S. Hussain, and R.M. Rehman, "Churn Analysis: Predicting Churns," IEEE International Conference on Digital Information Management (ICDIM), Phitsanulok, China, pp. 237-241, 2014.

[10] H.K. Qin, and J.H. Hu, “OLAP analysis of churn in the telecommunication's role,” IEEE International Conference on Computer Science and Service System (CSSS), Nanjing, China, pp. 2347-2350, 2011.

[11] X. Zhang, Z. Liu, X. Yang, W. Shi, and Q. Wang, “Predicting Customer Churn by Integrating the Effect of the Customer Contact Network,” IEEE International conference on Service Operations, Logistics and Informatics (SOLI), Shandong, China, pp. 392-397, 2010.

[12] Puspa, and G. Shobha, “Social Network classifier for churn prediction in Telecommunication Department,” IEEE International conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 1-7, 2013.

[13] S. Jumbo, L. Xim, and L. Wenhuang, “Application of AdaBoost in customer churn Prediction,” IEEE International Conference on Service, Systems and Service Management, Chengdu, China, pp. 1-6, 2007.

[14] J. Qi, Y. Zhang, and S. Shi, “TreeLogit model for customer churn prediction,” IEEE Asia-Pacific Conference on Service Computing, Gungdong, China, pp. 70-75, 2006.

[15] S. Salman, “Value based Time dimensioned Churn Prediction,” Journal of Emerging trends in Computing and Information, vol. 4, pp. 180-183, 2013.

[16] S- Y. Hung, D.C. Yen, and H. Wang, “Applying data mining to telecom churn management,” Expert Systems with Applications, vol. 31, No. 3, pp. 515–524, 2006.

[17] T. Mutanen, “Customers churn analysis – a case study,” Research Report No. VTT-R01184-06, Dated March 15, Available at: http://www.vtt.fi/inf/julkaisut/muut/2006/ customer_churn_case_study.pdf, Retrieved on 19 August 2008.

[18] S. A. Neslin, S. Gupta, W. Kamakura, J. Lu, and C. Mason, “Defection Detection: Improving Predictive Accuracy of Customer Churn Models,” Journal of Marketing Research, vol. 43, No. 2, pp. 204–211, 2006.

[19] B. H. Chu, M-S. Tsai, and C.S Ho, “Toward a Hybrid Data Mining Model for Customer Retention,” Knowledge-Based Systems, vol. 20, No. 8, pp. 703–718, 2007.

[20] M. V. Wezel, and R. Potharst, “Improved Customer Choice Predictions using Ensemble Methods,” European Journal of Operational Research, vol. 181, No. 1, pp. 436–452, 2007.

[21] D.A. Kumar, and V. Ravi, “Predicting Credit Card Customer Churn in Banks using Data Mining,” International Journal of Data Analysis Techniques and Strategies, vol. 1, No. 1, pp. 4–28, 2008.

[22] N. Naveen, V. Ravi, and D. Anil Kumar, “Application of fuzzyARTMAP for churn Prediction in Bank Credit Cards,” International Journal of Information and Decision Science, vol. 1, No. 4, pp. 428-444, 2009.

[23] Robert Hecht-Nielsen, “Counter propagation Networks,” Applied Optics, vol. 26, No. 23, pp. 4979-4984, 1987.

[24] G. A. Carpenter, and S. Grossberg, “Pattern Recognition by Self-Organizing Neural Networks,” Cambridge, MA: MIT Press, 1991.

[25] G. A. Carpenter, and S. Grossberg, “A self-organizing neural network for supervised learning, recognition and prediction,” IEEE Communication Magazine, vol. 30, No. 9, pp. 38–49, 1992.

[26] http://www.knime. Org/download.html

[27] J. R. Quinlan, “Decision trees as probabilistic classifiers,” In Proceedings of fourth international workshop machine learning, Irvine, CA, pp. 31–37, 1987.

[28] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, “Classification and Regression Trees,” Wadsworth, Belmont, CA, 1984.

[29] http:// www.salford-systems.com.

[30] T. Fawcett, “An introduction to ROC analysis,” Pattern Recognition Letters, vol. 27, pp. 861–874, 2006.