A PROCESS FRAMEWORK FOR SUBSCRIBER MANAGEMENT AND RETENTION IN NIGERIAN TELECOMMUNICATION INDUSTRY

J.O.Daramola, O.O. Oladipupo and A.G. Musa

Department of Computer and Information Sciences,Covenant University, Ota, Nigeria.

{wandi_ex, frajooje, bola_musa }@yahoo.com

Abstract

Churning which is a sudden defection of subscriber to competitors is a disturbing problem in the global telecommunication industry. Hence, a dominant approach for subscriber management and retention is churn control, since it is cheaper to retain an existing subscriber than acquiring a new one. Predictive modeling employs the use of data mining techniques to identify patterns and provide a result that a group of subscribers are likely to churn in the near future. However, the effectiveness of subscriber retention strategy in an organization can be further boosted if the reason for churn and the timing of churn can also be predicted.

In this paper, we propose a data mining process framework that can be used to predict churn, determine when a subscriber is likely to churn, provides the reason why a subscriber may churn, and recommend appropriate intervention strategy for customer retention using a combination of statistical and machine learning techniques. This experiment is carried out using data from a major telecom operator in Nigeria.

Key words: Churn prediction, telecommunication, decision support system, survival analysis, artificial neural networks.

1. Introduction

The competitive nature of the wireless telecommunication industry derives from having several wireless telecom operators striving to win over a particular population with services that are basically similar. This often leads to release of new services, promotion of incentives packages, endless advertisements and incessant new rate regimes. The effect of this, especially for a fully matured market is the tendency for a subscriber to migrate to another competitor. Although, the Nigerian wireless telecom market is not fully matured with less than 20% of the entire population as wireless subscribers, the elite population market is fully matured, and so churn control among this high-valued subscriber population should be of uppermost concern to any proactive telecom operator.

The phenomenon of churning is a process of defection of a subscriber from a company to its competitor. Churning is a major market repositioning situation which usually occurs when the market in which two or more companies compete, matures. When it comes, the whole organization has to change. It not only places a concern on a product, a person or a department but it is a challenge at all levels of the entire company hierarchy. To handle this, the company has to take a holistic view on the cause of the churning and what to do about it.

A general customer base analysis places an emphasis on effective churn management. Churning should be well understood by taking a cursory look at the phenomenon from the competition, the market and the external factor point of view. Of importance is the statistical ratio of customers loss to recently acquired customers. A high churn rate is a pointer to the fact that customers are dissatisfied with the quality of services rendered and a low or near competitor rate indicates the market is highly competitive and growing. When the rate is high, drastic measures should be taken to bring forth reasons for churning and strategies for subscriber retention.

Quite a number of models and techniques have been used to predict customer churn in literature. A large number have extremely delve into churn prediction and quite a number have extended the prediction to include when a customer is likely to churn and the reason(s) why a customer might churn [10],[11],[18]. Also, some of the techniques that have been used to predict customer churn includes: logistic regression [13], decision tree (Classification And Regression Tree - CART) [2], artificial neural network [3], [17], data mining [7],[14], boosting [13], and genetic algorithm [1].

In [13] Mozer C.M. et al. presented a scheme for predicting subscriber dissatisfaction and improving customer retention and profit maximization using statistical machine learning techniques, their work did not address the issues of why and when churn will likely take place. [5], is a case study report of churn prediction in telecommunication data using of a data mining software package called MiningMart. The focal point and critical success factor explored in this approach was a clever preprocessing of the given data. In [6] the capabilities of three predictive models at predicting churn namely: neural network, logistic regression and decision trees were compared using complaint data. The aim was to identify the most suitable model hence a statistical analysis was done on each of the models in terms of their predictive accuracy. In [8] was the report of a survey of churn and customer loyalty in the Korean telecom market.

Our approach in this work is not only to build a model that can predict which subscriber may churn, but with extended capabilities to advance reason(s) why subscriber may churn, when they may churn and recommend appropriate strategy for their retention. Therefore, our process framework handles churn prediction from four perspectives. We intend to implement this framework from a rich chunk of telecom subscribers' demographic data, subscribers' transactions information and subscribers' complaints information.

The outline of the other aspects of this paper is as follows: section 2 is a description of the process framework, in section 3 we gave a description of the experimental procedure with prototype data. In section 4 we discuss our result and in section 5 we have the conclusion.

2. The Process Framework

We discovered that many of the approaches reported in literature occur in contexts where the detailed information about subscriber like demography and transaction information are available, example is corporate clients. A dominant group of subscribers is the prepaid subscribers, where limited demographic information have been made available (an example is the Nigerian context), yet this group of subscribers forms the greater bulk of an operator's subscriber base. Our effort in this work is focused on building an appropriate prediction model from available information on prepaid subscribers. The process framework being proposed will facilitate data mining of telecommunication data such that it is possible to predict churn, adduce reasons for a possible churn and predict when churn will take place.

The figure below (see Figure 1) shows the framework architecture for subscriber management to facilitate churn prediction, predict when a customer is likely to churn, provide the likely reasons for a churn and give recommendations of intervention strategies for subscriber retention. Data from a telecom service provider about its subscribers comprising of the complaint and repair data, and customer transaction data will be used to build an appropriate data prediction model. The data model is then processed for churn prediction, and survival analysis. Thereafter, the results of the two analysis procedures are passed to a decision support expert system component that recommends the most appropriate intervention strategies for customer retention.

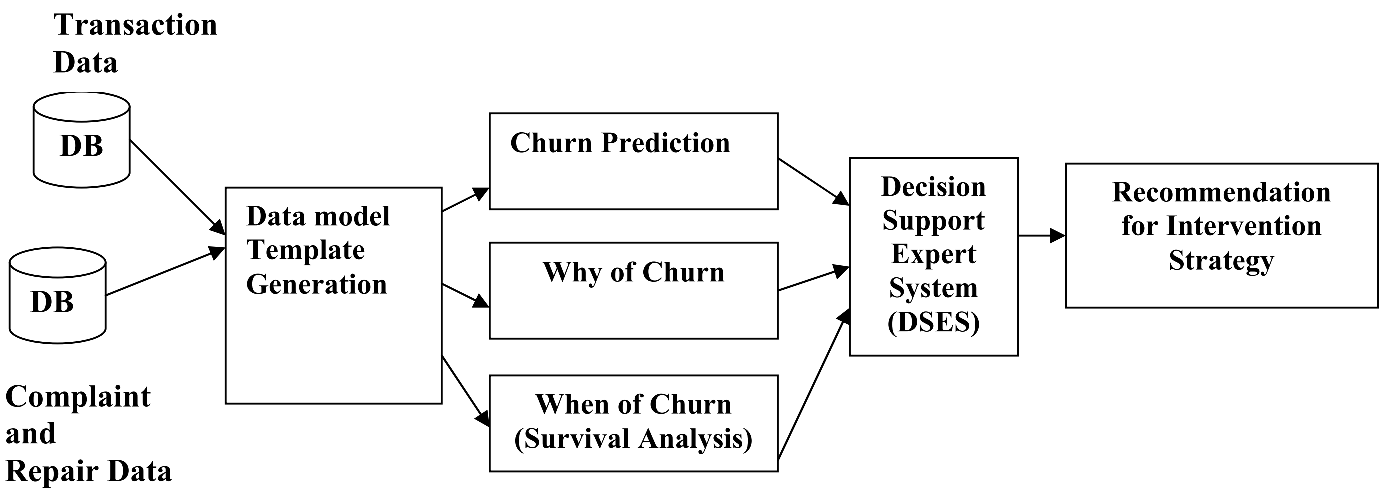


Figure 1: Framework Architecture for Subscriber Management and Retention

2.1 The input Data Set

The subscriber data used for our experiment was gleaned from the schema of a major wireless telecom operator (as the operator will not want to be identified, as churn prediction results are confidential). Although relatively few real data had been given, we were allowed access to the schema of the relevant databases. The schema of the subscriber transaction database is given as follows:

Subscriber = (*caller_number*, *called_number*, *incoming_route*, *outgoing_route*, *amount_b4_call*, *amount_after_call*, *International Mobile Subscriber Identity* (*IMSI)* (unique), *exchange_id*, *record_type* (sms, voice, gprs etc.), *event_type*). A typical subscribe transaction can be selected from the transaction database using appropriate query statement such as:

*"**Select** call date, details, caller_number, called_number, call_duration, call_charge, balance_before, balance_after **from** transactionDB **where** caller_number = subscriber_Number".*

The complaint data containing the records of complaint by subscribers has the structure:

Complaint_data = (*request_complaint_id* (unique), *date of complaint*, *time of complaint*, *type of complaint*, *status* (open or closed), *imputer* (internal staff initiator), *handled_by* (person to whom the problem was assigned)). The nature of the complaint database and subscriber transaction allows for multiple occurrences of a records of transaction and complaints of subscribers. Other relevant information about a subscriber include: *date of subscription, type service subscribed, subscriber number (unique)*

2.2 Data Prediction Model

The parameter template for predicting churn and the reason for churn were extracted from subscriber transactions and the complaint records datasets. A typical subscriber complaint falls into one of the established categories that have been denoted by unique codes, these are: swim swap billing (ssb), damaged recharged cards (drc), inabilibity to load credit (ilc), inability to make or receive calls (imr), roaming problems (rb), migration issues (mi), gprs (gp), wasp (wp), vas (vs) , dealers (dl), handset problems (hp), vip issues (vp) etc.

The parameters template as extracted from the input datasets used in our experiment is presented as follows:

1. Calling number (unique)
2. Calls made ratio = (total number of calls made / number of days)
3. Calls received ratio = (total number of calls received / number of days)
4. Calls dropped ratio = (total number of incomplete calls / number of days)
5. Amount charged ratio = (total amount charged / number of days)
6. Service utility ratio = (total calls duration / total number of calls)
7. Credit load ratio = (total credit load/ number of days)
8. ssb ratio = (number of ssb complaints/ total number of complaints)
9. drc ratio = (number of drc complaints/ total number of complaints)
10. ilc ratio = (number of ilc complaints/ total number of complaints)
11. imr ratio = (number of imr complaints/ total number of complaints)
12. rb ratio = (number of rb complaints/ total number of complaints)
13. gp ratio = (number of gp complaints/ total number of complaints)
14. wp ratio = (number of wp complaints/ total number of complaints)
15. vs ratio = (number of vs complaints/ total number of complaints)
16. dl ratio = (number of dl complaints/ total number of complaints)
17. hp ratio = (number of hp complaints/ total number of complaints)
18. vp ratio = (number of vp complaints/ total number of complaints)
19. Complaint frequency ratio = (number of complaint / number of days)
20. Response ratio = (number of closed complaints/ total number of complaints)
21. Mean-response = (total time spent on closed complaint / number of closed complaint)
22. No-response ratio = (total number of open complaint / total number of complaints)

3. The Experiment

3.1 Churn Prediction with ANN

We implemented an algorithm to scan the input datasets (i.e. transaction and complaint data) in order to generate the values for the parameters of the prediction template. Table 1 shows sample data values after the generation process. The values in the prediction data model were rescaled using interval scaling with the formula:

$$\mathbf{A_i = (V_i - V_{min}) / (V_{max} - V_{min}) \quad For\ i = 1 \ldots n} \qquad (1)$$

The rescaled values were used as inputs into a feed-forward back propagation artificial neural network model (see Table 2). The network outputs are also unscaled using the formula:

$$\mathbf{V_i = A_i (V_{max} - V_{min}) + V_{min} \quad For\ i = 1 \ldots n} \qquad (2)$$

The churn prediction neural network model was implemented using MATLAB 7.0 Software [12]. The configuration of the feedforward neural network model used was a 21-22-1 Multi-Layer Perceptron (MLP) [4], with 21 neuronode units in the input layer which directly corresponds to the number of input parameters in our prediction model template, 22 neuronode units in the hidden layer (1) and one neuronode unit in the output layer that corresponds to the output (1 or 0) which is indicative of whether a subscriber churns or not. The Levenberg-Marquardt (TRAINLM) backpropagation algorithm was used in training the network with the following training parameter values:

net.trainParam.epochs 500	*'Maximum number of epochs to train*
net.trainParam.goal 0	*'Performance goal*
net.trainParam.max_fail 5	*'Maximum validation failures*
net.trainParam.mem_reduc 1	*'Factor to use for memory/speed tradeoff*
net.trainParam.min_grad 1e-010	*' Minimum performance gradient*
net.trainParam.mu 0.001	*' Initial Adaptive learning parameter value (Mu)*
net.trainParam.mu_dec 0.1	*'Mu decrease factor*
net.trainParam.mu_inc 10	*'Mu increase factor*
net.trainParam.mu_max 1e12	*'Maximum Mu*
net.trainParam.show 25	*'Epochs between showing progress*
net.trainParam.time inf	*' interval*

Figure 2 show the graphical representation of pattern of weight convergence and number of training epochs.

080x2	3500/182	2650/182	56/ 182	75400 /182	15600/ 3120	56800/ 182	65/ 700	160/ 700	80/ 700	75/ 700	83/ 700	45/ 700	90/ 700	40/ 700	15/ 700	39/ 700	8/ 700	700/ 182	600/ 700	2006/ 600	100/700	1
080x3	2560/100	160/ 100	30/ 100	89000/ 100	10000/ 2720	6800/ 100	40/ 200	10/ 200	30/ 200	15/ 200	8/200	7/ 200	2/ 200	50/ 200	20/ 200	10/ 200	8/ 200	65/100	63/65	900/63	2/65	0
080x4	5000/365	2010/365	20/ 365	125005/365	18955/ 7010	120000/365	150/ 850	72/ 850	45/ 850	99/ 850	80/ 850	40/ 850	65/ 850	100/ 850	40/ 850	9/850	150/850	950/ 365	950/ 950	1960/ 950	100/950	0
080x5	4560/270	4000/270	10/ 270	56892/ 270	10690/ 8560	50565/ 270	100/ 1200	110/ 1200	70/ 1200	5/ 1200	280/ 1200	63/ 1200	79/ 1200	95/ 1200	106/ 1200	215/ 1200	77/ 1200	550/ 270	450/ 550	4698/ 450	100/550	1

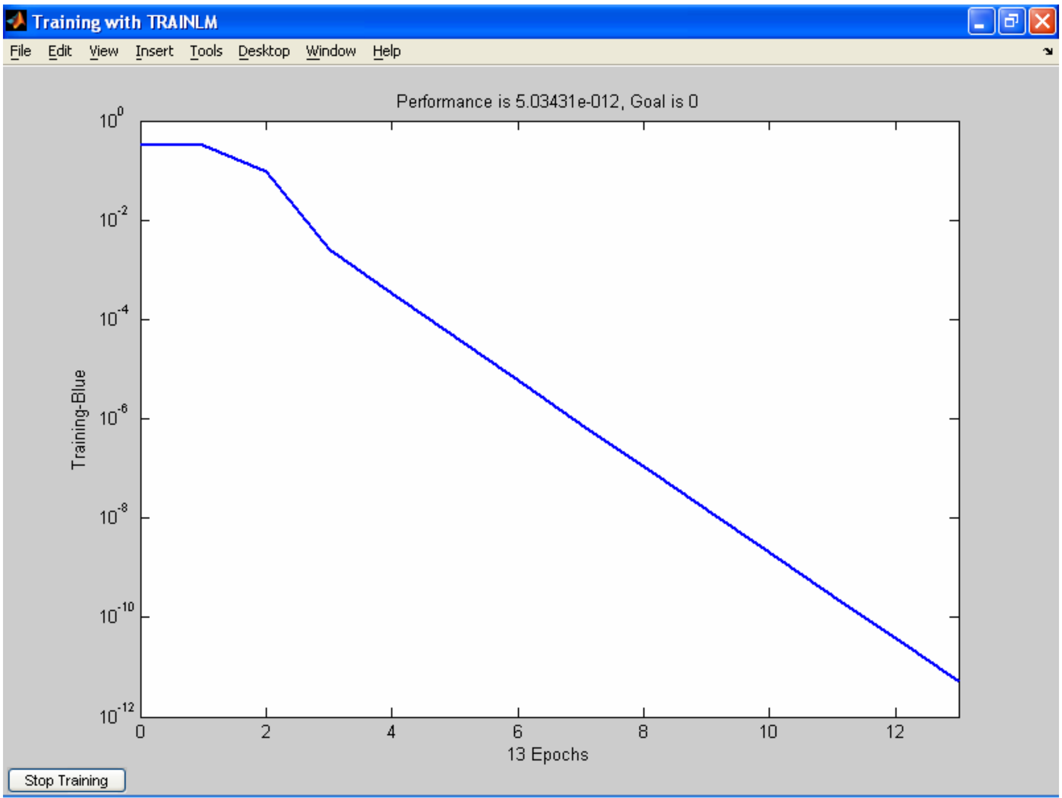
Table 2: Showing ratio values from subscriber data

Calling number	Calls made ratio	Calls received ratio	Calls dropped ratio	Amount charged ratio	Service utility ratio	Credit load ratio	ssb ratio	drc ratio	ilc ratio	imr ratio	rb ratio	gp ratio	wp ratio	vs ratio	dl ratio	hp ratio	vp ratio	Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
080x1	17.00	13.72	0.29	429.88	2.36	270.41	0.15	0.05	0.38	0.24	0.03	0.07	0.09	0.13	0.01	0.07	0.06	2.88	0.96	4.96	0.04	1
080x2	19.23	14.56	0.31	414.29	5.00	312.09	0.09	0.23	0.11	0.11	0.12	0.06	0.13	0.06	0.02	0.06	0.01	3.85	0.86	3.34	0.14	1
080x3	25.60	1.60	0.30	890.00	3.68	68.00	0.20	0.05	0.15	0.08	0.04	0.04	0.01	0.25	0.10	0.05	0.04	0.65	0.97	14.29	0.11	0
080x4	13.70	5.51	0.05	342.48	2.70	328.77	0.18	0.08	0.05	0.12	0.09	0.05	0.08	0.12	0.05	0.01	0.18	2.60	1.00	2.06	0.11	0
080x5	16.89	14.81	0.04	210.71	1.25	187.28	0.08	0.09	0.06	0.00	0.23	0.05	0.07	0.08	0.09	0.18	0.06	2.04	0.82	10.44	0.18	1

Table 3: Showing rescaled values for neural network input (split into two halves)

Calling number	Calls made ratio	Calls received ratio	Calls dropped ratio	Amount charged ratio	Service utility ratio	Credit load ratio	ssb ratio	drc ratio	ilc ratio	imr ratio	rb ratio
080x1132110	0.019101	0.015410	0.000323	0.483011	0.002657	0.303833	0.000166	0.000060	0.000426	0.000268	0.000039
080x1332710	0.021608	0.016360	0.000346	0.465490	0.005618	0.350661	0.000104	0.000257	0.000128	0.000120	0.000133
080x5612770	0.028764	0.001798	0.000337	1.000000	0.004131	0.076404	0.000225	0.000056	0.000169	0.000084	0.000045
080x3434160	0.015392	0.006187	0.000062	0.384808	0.003038	0.369401	0.000198	0.000095	0.000059	0.000131	0.000106
080x1752819	0.018976	0.016646	0.000042	0.236754	0.001403	0.210424	0.000094	0.000103	0.000066	0.000005	0.000262

gp ratio	wp ratio	vs ratio	dl ratio	hp ratio	vp ratio	Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
0.000074	0.000106	0.000147	0.000011	0.000080	0.000070	0.003232	0.001081	0.005568	0.000043	1
0.000072	0.000144	0.000064	0.000024	0.000063	0.000013	0.004322	0.000963	0.003757	0.000161	1
0.000039	0.000011	0.000281	0.000112	0.000056	0.000045	0.000730	0.001089	0.016051	0.000118	0
0.000053	0.000086	0.000132	0.000053	0.000012	0.000198	0.002924	0.001124	0.002318	0.000118	0
0.000059	0.000074	0.000089	0.000099	0.000201	0.000072	0.002289	0.000919	0.011730	0.000204	1



3.2 Survival Data mining

The survival time analysis of subscriber data to estimate the life time value of a subscriber and the churn hazard per subscriber over a time period will be carried out using SAS software [15],[16]. Survival data mining describes the distribution of the survival time for customers in a given population, investigates the strength of parameter influence on expected survival time and allows comparing of survival time distributions among different subpopulations [9]. This analysis gives insight into customers' behaviours and finds ways to increase their survival time. This phase of the experiment has not been concluded, it is intended to determine the survival rate of each subscriber in the population and predict at what point a subscriber is likely to churn.

3.3 Generating Reasons for Churn and Intervention Strategy

The results obtained from churn prediction using artificial neural network and churn survival analysis will be fed into a Decision Support Expert System (DSES) which will generate the probable reasons for churn and recommendations of intervention strategies for customer retention using a knowledge rule-based inference engine. The inference engine of the DSES is composed of a set of if - then rules that provides recommendations of appropriate incentives based on the credit rating of a subscriber. Subscribers are classified as high-valued, medium-valued and low-valued using a specific value rating threshold function Th(x) given as:

$$\mathbf{Th(x)} = \begin{cases} 1 & \text{if } \mathbf{V(x)} \geq 60 \quad \text{// 'high} \\ 0.5 & \text{if } \mathbf{V(x)} \geq 40 \text{ and } \mathbf{V(x)} < 60 \quad \text{// 'medium} \\ 0 & \text{if } \mathbf{V(x)} < 40 \text{ // 'low} \end{cases} \qquad (3)$$

Where:

V(x) = ((calls made ratio + calls received ratio) / 2 * 100) (4)

Generally, low valued customers will be ignored while medium and high valued customers will have best-fit retention strategies recommended for them using a set of in-built rules.

In other to generate reasons for churn, the DSES reckons a particular complaint ratio metric as 'high' if it is above 0.5 and 'low' otherwise. Typical instances of inference rules for generating reasons for churn are shown as follows:

1. If call dropped ratio $\geq$ 0.5 ('high') and churn prediction = 1 ('Yes') Then 'subscriber may churn due to high dropped calls'

2. if ssb ratio $\geq$ 0.5 and call dropped ratio $\geq$ 0.5 and churn prediction = 1 Then ' subscriber may churn due to ssb problems, high dropped calls'

3. if ilc ratio $\geq$ 0.5 and ssb ratio $\geq$ 0.5 and call dropped ratio $\geq$ 0.5 and churn prediction = 1 Then 'subscriber may churn due to ilc, ssb problems, high dropped calls'

4. Results

Thus far, we have completed the implementation of the neural network-based churn prediction and DSES components of the architectural framework, and the results obtained have been very encouraging. Although more time is still needed to conclude work on the aspect of survival analysis, our experience so far gives credence to the feasibility of this process framework as a viable model for effective subscriber management in the telecommunication industry. The DSES proves to be an ideal complement to the interpolative power of the neural network in predicting prospective future churn by providing the probable reasons why churn may occur and a suggestion of possible intervention strategies. All of this has so far been achieved with an acceptable level of accuracy.

5. Conclusion

A comprehensive process framework for management of churn that embraces churn prediction, determination of why and when of churn and automatic recommendation of retention strategies like the one described in this work offers a viable platform for effective subscriber management in any fast maturing telecom market. This kind of holistic approach that will ensure the retention of high-valued customers and ultimately the promotion of profitability is therefore recommended. We intend to complete the other aspects of this work in order to further drive home the novelty of our approach in contrast to many other existing approaches.

References

1. Au, W., Chan, C. C. and Yao, X. (2003): A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction, IEEE transactions on evolutionary computation, 7, 6, 532-545.
2. Bloemer, J., Brijis, T., Vanhoof, K. and Swinnen, G. (2002): Comparing Complete and Partial Classification for Identifying Customers at Risk, International journal of research in marketing, 20, 117-131.
3. Boone, D. S. and Roehm, M. (2002): Retail Segmentation Using Artificial Neural Networks, International journal of research in marketing, 19, 287-301.
4. Daramola, J., Olugbara, O., Owoloko, E. and Musa, A (2006): Decision Support for Property Appraisal Using a Neural-CBR Intelligent System., Proceedings of the International Conference on New Trends in Mathematical and Computer Sciences With Applications to Real World Problems (NTMCS), Covenant University, Ota, Nigeria [65-82].

5. Euler, T. (2005): Churn Prediction in Telecommunication Using MiningMart, In: Proceedings of the Workshop on Data Mining and Business (DMBiz) at the 9th European Conference on Principles and Practice in Knowledge Discovery in Databases (PKDD), Porto, Portugal.
6. Hadden, J. Tiwari, A., Roy, R, and Ruta, D (2006): Churn Prediction using Complaints Data, Transactions on Engineering, Computing and Technology, Enformatika, Vol. 13, ISBN 1305-5315, pp 158-163.
7. Hwang, H., Taesoo, J. and Euiiho, S. (2004): An Ltv Model and Customer Segmentation Based on Customer Value: A Case Study on the Wireless Telecommunications Industry, Expert systems with applications, 26, 181-188.
8. Kim, H. and Yoon, C. (2004): Determinants of Subscriber Churn and Customer Loyalty in the Korean Mobile Telephony Market, Telecommunications Policy, 28, 751-765.
9. Linoff, G. (2004) : Survival Data Mining for Customer Insight, Data miners, http://www.data-miners.com
10. Lu, J (2001): Detecting Churn Triggers for Early Life customers in the Telecommunication Industry- An application of Interactive Tree Training, Proceedings of the 2nd Data mining Conference of DiaMondSUG 2001, Chicago IL.
11. Lu, J. (2002): Predicting Customer Churn in the Telecommunications Industry- An Application of Survival Analysis Using SAS, Proceedings of SUG127, Orlando, Florida, Paper 114-27.
12. Mathworks Inc. (2003): MATLAB © Software, version 7.0
13. Mozer, M., Wolniewicz, R., Grimes, D, Johnson, E., and Kaushansky, H (2000): Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunication Industry, IEEE Transactions on Neural Networks, Special Issue on Data Mining and knowledge Representation
14. Rygielski, C., Wang, J. and Yen, D. C. (2002): Data Mining Techniques for Customer Relationship Management, Technology in Society, 24, 483-502.
15. SAS Institute Inc. (1989), SAS/STAT @ Software: Changes and Enhancements through Release 6.11, Volume 2, Cary, NC: SAS Institute Inc. 846pp.
16. SAS Institute Inc. (1989), SAS/STAT @ Users Guide, Version 6.0, Fourth Edition, Volume 2, Cary, NC: SAS Institute Inc. 846pp.
17. Vellido, A., Lisboa, P. J. G. and Meehan, K. (1999): Segmentation of the Online Shopping Market Using Neural Networks, Expert systems with applications, 17, 303-314.
18. Wei, C. and Chiu, I. (2002): Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach, Expert systems with applications, 23, 103-112.