IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

A Data Mining Process Framework for Churn Management in Mobile Telecommunication Industry

Daramola, J. Olawande, Oladipupo, O. Olufunke, Musa, G. Adebola
Department of Computer and Information Sciences
Covenant University, Ota, Nigeria.
dwande@gmail.com, frajooje@yahoo.com, bola_musa@yahoo.com

ABSTRACT

Churning which is the sudden defection of a subscriber to competitors is a disturbing problem in the global telecommunication industry. However, the effectiveness of existing churn control strategies can be improved if an integrated approach that incorporates several dimensions of the phenomenon of churning is adopted. In contrast to exiting approaches, this paper proposes an integrated approach to churn management and control by using a data mining process framework that enables churn prediction, determination of reason(s) for churn, and recommendation of appropriate intervention strategy for customer retention. A data mining experiment that was undertaken using data from a major telecom operator in Nigeria to assess the viability of the approach yielded encouraging results.

Key words: Churn prediction, subscriber retention, mobile telecommunication, decision support system, artificial neural networks.

1. Introduction

The competitive nature of the wireless telecommunication industry derives from having several wireless telecoms 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 40% of the entire population as wireless subscribers, the elite market is fully saturated, and so churn control among this high-valued subscriber population should be of uppermost concern to any telecom operator that wants to be proactive.

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 occur as the market in which two or more companies are competing matures. When this is the case, 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 place emphasis on effective churn management. Churning should be well understood by taking a thorough look at the phenomenon from the competition, the market and the external factors point of view. The statistical ratio of customer's loss to recently acquired customers is also an important metric for determining churn. 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.

Different kinds 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 few have extended the prediction to include when a customer is likely to churn and the reason(s) why a customer might churn [1,2,3]. Also, some of the techniques that have been used to predict customer churn includes: logistic regression [4], decision tree (Classification And Regression Tree - CART) [5], artificial neural network [6], [7], data mining [8],[9], boosting [4], genetic algorithm [10] and ordinal regression [11].

In [4], Mozer et al. presented a scheme for predicting subscriber dissatisfaction and improving customer retention and profit maximization using statistical machine learning techniques, but the work did not address the issues of why and when churn will likely take place. In [11], a model based on the use of ordinal regression was adopted as an alternative to survival analysis in determining the tenure of customer in the mobile telecommunication industry. In [12], is a case study report of churn prediction in telecommunication industry using a data mining software package called *MiningMart*. The critical success factor of approach used in the work was a clever preprocessing of the given data. In [13], 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 [14], the report of a survey of churn and customer loyalty in the Korean telecom market was given.

However, to the best of our knowledge, none of the existing churn prediction approaches in the telecom industry has adopted an integrated approach to churn management and control that incorporates several dimensions of the

IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

phenomenon of churning in a single process framework. In this work, a data mining process framework that enables churn prediction, determination of reasons for churn, and recommendation of appropriate intervention strategy for customer retention is presented. By using a combination of expert systems and machine learning techniques, the process framework handles churn prediction from three perspectives. These are 1) prediction of which subscriber may churn; 2) Determination of reason(s) why subscriber may churn; and 3) recommendation of appropriate strategy for customer retention. The process framework was implemented using 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 using prototypical data. In section 4 we discuss our results and in section 5 we give the conclusion.

2. Description of the Process Framework

Most of the approaches so far reported in literature occur in contexts where the detailed information about subscribers like demography and transaction information are available; an example is corporate clients [12,13,14]. 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 form the greater bulk of an operator's subscriber base. The effort in this work is focused on building an appropriate prediction model from available information on prepaid subscribers. This prediction model provides the basis for a process framework that facilitates the mining of telecommunication data such that the possibility of churn can be predicted; the reasons for churn can be provided; and appropriate remedy recommended.

The proposed framework architecture for subscriber management that facilitates churn prediction, discovery of reasons for churn, and recommendations of intervention strategy for subscriber retention is shown in figure 1. It is composed of subscribers' data components as obtained from a telecommunication service provider. The subscribers' data comprises of the complaint and repair data, and customer transaction data which are used to build an appropriate data prediction model. The data prediction model is processed for churn prediction. Thereafter, the result of the churn prediction procedure is passed to a decision support expert system component that recommends the most appropriate intervention strategy for customer retention.

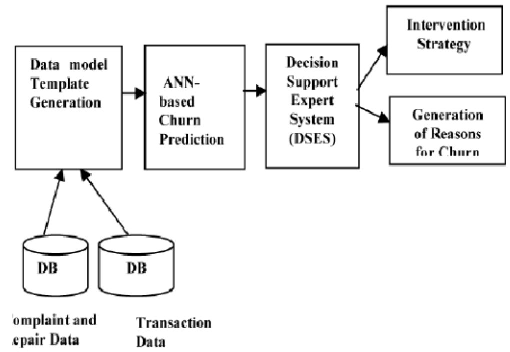

Figure 1: Framework Architecture for Subscriber Management and Retention

3. The Experiment

A data mining experiment was carried out using data obtained from the telecomm industry in order to validate the plausibility of our proposed framework architecture for churn management. The subscriber data used for the experiment was obtained from the schema of a major wireless Telecoms Operator in Nigeria (the specific operator will not want to be identified, as churn prediction results are confidential). Although relatively few real data was given out, the schema of the relevant database tables were made accessible from which further hypothesized data were generated for the purpose of the experiment. The outlook of the schema of the subscriber transaction database is 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). Therefore a typical subscriber 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 records of transactions and complaints of subscribers. Other relevant information about a subscriber include: date of subscription, type of service subscribed, and subscriber number (unique).

IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

3.1 Data Prediction Model

In order to predict churn and generate the reason for churn, a prediction model that exploits the observable feature characteristics of individual subscriber transactions in the database of the Telecoms Operator was formulated. The structure of the complaints dataset of the Telecoms Operator revealed that a typical subscriber complaint can be classified into unique categories that have been denoted by unique codes. These include : 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), and vip issues (vp). Twenty one (21) metric parameters that were considered most significant in determining the churn behavior of a typical subscriber were identified from the subscriber transactions and the complaint records datasets. Data for these metric parameters were extracted from the subscriber transactions and the complaint records datasets in order to obtain a data model that can be used for churn prediction. An algorithm was implemented to scan the input datasets (i.e. transaction and complaint data) in order to generate the values for the metric parameters per subscriber. Table 1 shows sample data values after the generation process. The generated data model consist of 1,250 subscriber records, which was partitioned into training dataset (80%) and test dataset (20%). The training dataset of the data model was used for the Artificial Neural Network-based prediction of churn. The 21 metric parameters are 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)

3.2 Churn Prediction with Artificial Neural Network (ANN)

The values in the data model were rescaled to values between 0 - 1 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 churn prediction neural network model was implemented using MATLAB 7.0 Software [15]. The configuration of the feedforward neural network model used was a 21-22-1 Multi-Layer Perceptron (MLP) [16], 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 (only one) 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_reduc1 *'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*

In figure 2 the graphical representation of pattern of weight convergence and number of training epochs is presented.

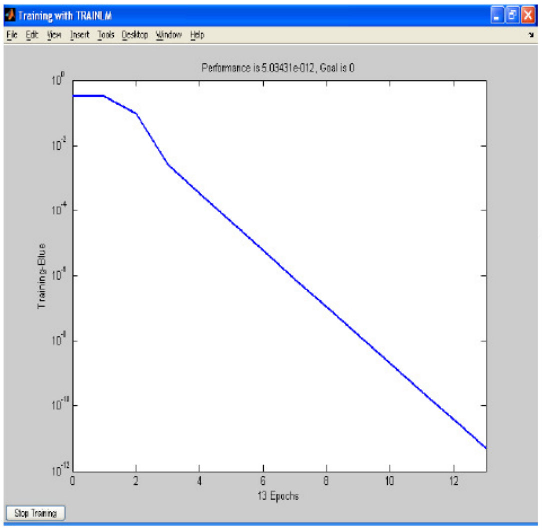

Figure 2: Graph of Weight Convergence vs. Number of Epochs

3.3 Generating Reasons for Churn and Intervention Strategy

The results obtained from churn prediction using artificial neural network was passed into a Decision Support Expert System (DSES) to generate the probable reasons for churn, and recommendations of intervention strategy for customer retention. The DSES was implemented using the Java Expert System Shell (JESS) 7.1, a rule engine for the Java Platform [17]. The rule inference engine of the DSES was populated with a set of if - then rules that enabled the generation of 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:

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

Where: $\mathbf{V(x)}$ = ((calls made ratio + calls received ratio) / 2 * 100) (4)

In the recommendation scheme, the low valued customers are ignored while medium and high valued customers have best-fit retention strategies recommended for them based on the 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 ≥ 0.5 ('high') and churn prediction = 1 ('Yes') Then
'subscriber may churn due to high dropped calls'
2. if ssb ratio ≥ 0.5 and call dropped ratio ≥ 0.5 and churn prediction = 1 Then
' subscriber may churn due to ssb problems, high dropped calls'
3. if ilc ratio ≥ 0.5 and ssb ratio ≥ 0.5 and call dropped ratio ≥ 0.5 and churn prediction = 1 Then 'subscriber may churn due to ilc, ssb problems, high dropped calls'

The equivalent Jess definition of these rules is shown in figure 3, while figure 4 is a snapshot of the churn management and control interface of the DSES.

```
(defrule high-dropped-calls; checks what happens for
( >= ?call_dropped_ratio 0.5) &: (eq ?churn_prediction 1)
=> ( Printout t "subscriber may churn due to high dropped
calls" clf)
)
(defrule ssb-problems; checks what happens for
( >= ? ssb_ratio 0.5) &: (>= ?call_dropped_ratio 0.5)
&: (eq ?churn_prediction 1)
=> ( Printout t "subscriber may churn due to ssb, high
dropped calls" clf)
)
(defrule ilc_ratio-ssb-high-dropped-calls; checks what
happens for
( >= ? ilc_ratio 0.5) &: ( >= : ssb_ratio 0.5)
&: (>= ?call_dropped_ratio 0.5) &: (eq ?churn_prediction 1)
=> ( Printout t "subscriber may churn due to ilc, ssb
problems , high dropped calls" clf)
)
```

Figure 3: A Sample of Jess Rules in the DSES

4. Results and Discussions

Thus far, our experience gives credence to the feasibility of using the proposed 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 strategy. All of these have been achieved with an acceptable level of performance. Also, it was discovered that using a single integrated framework that incorporated several dimensions of the churn phenomenon enabled information about subscribers that are useful for churn control to be effectively managed. The framework made it possible to detect likely churn early enough and to devise appropriate retention strategy to forestall churn. This offers significant improvement over existing churn control approaches that lack this kind of multi-dimensional orientation.

5. Conclusion

In this paper, an integrated data mining process framework for management of churn that enables churn prediction, determination of reasons for churn and recommendation of appropriate retention strategy for

IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

effective subscriber management in a fast maturing telecom market has been proposed. The viability of the concept proposed was confirmed by conducting a data mining experiment using data from the Nigerian telecom industry. The results from the experiment revealed that the approach have the potential to ensure the retention of high-valued customers and ultimately the promotion of profitability. In future work, the data mining framework will be extended to include the detection of the timing of churn i.e. (when a subscriber is likely to churn) using survival analysis.

References

[1] 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, Chicago IL.

[2] 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.

[3] Wei, C. and Chiu, I., 2002: Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach, Expert systems with applications, 23, 103-112.

[4] 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

[5] 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.

[6] Boone, D. S. and Roehm, M. ,2002: Retail Segmentation Using Artificial Neural Networks, International journal of research in marketing, 19, 287-301.

[7] 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.

[8] 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.

[9] Rygielski, C., Wang, J. and Yen, D. C., 2002: Data Mining Techniques for Customer Relationship Management, Technology in Society, 24, 483-502.

[10] 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.

[11] Gopal, R.K. and Meher, S.K., 2008: Customer Churn Time Prediction in Mobile Telecommunication Industry Using Ordinal Regression, Lecture Notes in Computer Science, Volume 5012, Springer Berlin / Heidelberg

[12] 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.

[13] 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.

[14] Kim, H. and Yoon, C., 2004: Determinants of Subscriber Churn and Customer Loyalty in the Korean Mobile Telephony Market, Telecommunications Policy, 28, 751-765.

[15] Mathworks Inc., 2003: MATLAB © Software, version 7.0

[16] 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, pp 65-82.

[17] JESS: Java Expert System Rule Engine 7.1, Available at (http// herzberg.ca.sandia.gov/jess/docs/index.shtml)

IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

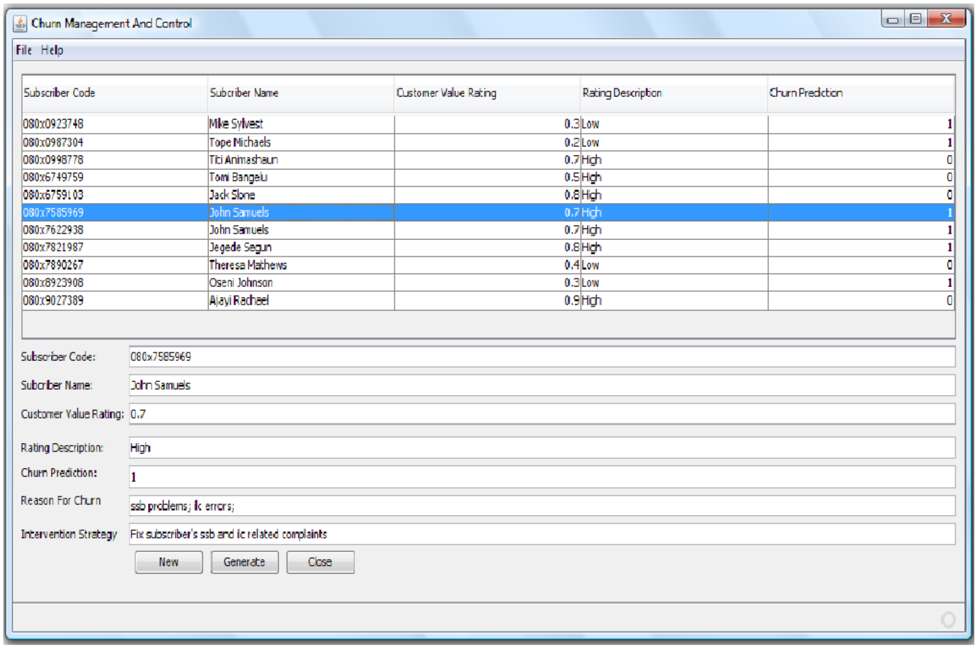

Figure 4: A view of the Churn Management and Control Interface

Table 1: Showing values generated from subscriber data for Churn Prediction

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	Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
080x 1	6205/365	5006/365	105/365	156906/365	26509/11211	98700/365	155/1050	56/1050	98/1050	1050/365	1010/1050	5005/1010	40/1050	1
080x 2	3500/182	2650/182	56/182	75400/182	15600/3120	56800/182	65/700	160/700	80/700	700/182	600/700	2006/600	100/700	1
080x 3	2560/100	160/100	30/100	89000/100	10000/2720	6800/100	40/200	10/200	30/200	65/100	63/65	900/63	2/65	0
080x 4	5000/365	2010/365	20/365	125005/365	18955/7010	120000/365	150/850	72/850	45/850	950/365	950/950	1960/950	100/950	0
080x 5	4560/270	4000/270	10/270	56892/270	10690/8560	50565/270	100/1200	110/1200	70/1200	550/270	450/550	4698/450	100/550	1

IJCSS, Vol.1, No.1, 2009
ISSN: 1803-8336

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	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	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	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	0.65	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	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	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
080x1132110	0.019101	0.015410	0.000323	0.483011	0.002657	0.303833	0.000166	0.000060	0.000426
080x1332710	0.021608	0.016360	0.000346	0.465490	0.005618	0.350661	0.000104	0.000257	0.000128
080x5612770	0.028764	0.001798	0.000337	1.000000	0.004131	0.076404	0.000225	0.000056	0.000169
080x3434160	0.015392	0.006187	0.000062	0.384808	0.003038	0.369401	0.000198	0.000095	0.000059
080x1752819	0.018976	0.016646	0.000042	0.236754	0.001403	0.210424	0.000094	0.000103	0.000066

Complaint frequency ratio	Response ratio	Mean-response	No-response ratio	Churn
0.003232	0.001081	0.005568	0.000043	1
0.004322	0.000963	0.003757	0.000161	1
0.000730	0.001089	0.016051	0.000118	0
0.002924	0.001124	0.002318	0.000118	0
0.002289	0.000919	0.011730	0.000204	1