Computing, Information Systems
& Development Informatics Journal

Volume 3. No. 2. May, 2012

Data Mining Technique for Predicting Telecommunications Industry Customer Churn Using both Descriptive and Predictive Algorithms

Kolajo Taiwo
Computer Science Department
Federal College of Education (Technical)
Bichi, Kano, Nigeria
taiwo_kolajo@yahoo.com

Adeyemo, A.B.
Computer Science Department
University of Ibadan,
Ibadan, Nigeria
sesan_adeyemo@yahoo.com

Reference Format: Kolajo, T & Adeyemo, A.B. (2012). Data Mining Technique for Predicting Telecommunications Industry Customer Churn Using both Descriptive and Predictive Algorithms. Computing, Information Systems & Development Informatics Journal. Vol 3, No.2. pp 27-34

Data Mining Technique for Predicting Telecommunications Industry Customer Churn Using both Descriptive and Predictive Algorithms

Kolajo, T & Adeyemo, A.B

ABSTRACT
As markets have become increasingly saturated, companies have acknowledged that their business strategies need to focus on identifying those customers who are most likely to churn. It is becoming common knowledge in business, that retaining existing customers is the best core marketing strategy to survive in industry. In this research, both descriptive and predictive data mining techniques were used to determine the calling behaviour of subscribers and to recognise subscribers with high probability of churn in a telecommunications company subscriber database. First a data model for the input data variables obtained from the subscriber database was developed. Then Simple K-Means and Expected Maximization (EM) clustering algorithms were used for the clustering stage, while Decision Stump, M5P and RepTree Decision Tree algorithms were used for the classification stage. The best algorithms in both the clustering and classification stages were used for the prediction process where customers that were likely to churn were identified.

Keywords: customer churn; prediction; clustering; classification

1. INTRODUCTION

The mobile telephony market is one of the fastest growing service segments in telecommunications, and more than 75% of all potential phone calls worldwide can be made through mobile phones and as with the any other competitive markets, the mode of competition has shifted from acquisition to retention of customers [1].

Among all industries that suffer customer churn the telecommunications industry can be considered as being at the top of the list with an approximate annual rate of 30% [2]; [3]. This results in a waste of money and effort and "is like adding water to a leaking bucket" [4]. Considering the fact that the cost to European and US telecommunication companies is US$ 4 billion per year, then it seems reasonable to invest more on churn management rather than acquisition management for mature companies, especially when it is noted that the cost of acquiring new customer is eight times more than that of retaining an existing one [3]. On the other hand, existing subscribers tend to generate more cash flow and profit, since they are less sensitive to price and often lead to sales referrals [5]. Due to the high cost of acquiring new subscribers and considerable benefits of retaining the existing ones, building a churn prediction model to facilitate subsequent churn management and customer retention is critical for the success or bottom-line survival of a mobile telecommunications provider in this greatly compressed market-space.

Subscriber churning (often referred to as customer attrition in other industries) in mobile telecommunication refers to the movement of subscribers from one provider to another. Many subscribers frequently churn from one provider to another in search of better rates or services. Churning customers can be divided into two main groups, voluntary churners and non-voluntary churners. Non-voluntary churn is the type of churn in which the service is purposely withdrawn by the company. Voluntary churn is more difficult to determine. This type of churn occurs when a customer makes a conscious decision to terminate his/her service with the provider. This type of churn has been a serious and puzzling problem for service providers.

The varied behaviour of consumers has baffled researchers and market practitioner's alike [6]. Voluntary churn can be divided into two sub categories, incidental churn and deliberate churn. Incidental churn happens when changes in circumstances prevent the customer from further requiring the provided service and is a small percentage of a company's voluntary churn [7]. Deliberate churn is the problem that most churn management solutions attempt to identify. This type of churn occurs when a customer decides to move to a competing company due to reasons of dissatisfaction [1]. Deliberate churn within the telecommunications industry is minimised because switching would require a change in the telephone number. In 2003 customers in the United States of America were given the option to switch mobile telephone provider but keep their existing phone number, and as soon as this law came into force 12 million customers immediately churned from their service providers thereby increasing the retention battle [8].

A churn management solution should not target the entire customer base because (i) not all customers are worth retaining, and (ii) customer retention costs money; attempting to retain customers that have no intention of churning is a waste of resources. Nowadays lack of data is no longer a problem, but the inability to extract useful information from data [9]. Due to the constant increase in the amount of data efficiently operable to managers and policy makers through the high speed computers and rapid data communication, there has grown and will continue to grow a greater dependency on statistical methods as a means of extracting useful information from the abundant data sources. To survive or maintain an advantage in an ever-increasing competitive marketplace, many companies are turning to data mining techniques to address churn prediction and management [10].

Data mining is the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner [11]. Data mining is an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases, and visualization to address the issue of information extraction from large data bases. Data mining techniques include a wide range of choices from many disciplines. These choices include techniques such as support vector machines, correlation, linear regression, nonlinear regression, genetic algorithms, neural networks, and decision trees. The choice of a data mining technique is contingent upon the nature of the problem to be solved and the size of the database. Based on the kind of knowledge which can be discovered from databases, data mining techniques can be broadly classified into several categories including clustering, classification, dependency analysis, data visualization and text mining [12].

Clustering analysis is a process whereby a set of instances (without a predefined class attribute) is partitioned (or grouped) according to some distance metric into several clusters in which all instances in one cluster are similar to each other and different from the instances of other clusters. Classification is a model that induces a model to categorize a set of pre-classified instances (called training examples) into classes. Such a classification model is now used to classify future instances. Clustering is a way to segment data into groups that are not previously defined, whereas classification is a way to segment data by assigning it to groups that are already defined. Dependency analysis discovers dependency patterns (e.g. association rules, sequential patterns, temporal patterns, and episode rules) embedded in data. Data Visualization allows decision makers to view complex patterns in the data as visual objects in three dimensions and colour; it supports advanced manipulation capabilities to slice, rotate or zoom the objects to provide varying levels of details of the patterns observed.

In this study both descriptive model and predictive data mining techniques will be used to extract information on the calling behaviour of subscribers and to recognise subscribers with high probability of churn in the future. While some researchers have focused on the use of either descriptive or descriptive algorithms, in this work both algorithms will be combined. First the elements of the dataset will be grouped by clustering them and then classification algorithms will be applied to the clusters of interest so that each cluster's unique "rules" for relating attributes to classes can be determined and thereby more accurately classify the members of each cluster. The dataset used were obtained from a Nigerian Telecommunications service provider.

2. MATERIALS AND METHODS

The Customer churn prediction model was developed based on some selected input variables from a Nigerian telecommunications service provider customer database.

2.1 Data Selection and Preprocessing

The data used were from the call records of subscribers in one of the telecommunications service providers in Nigeria. The total number of records in the dataset is 228,520. The records were for transactions covering a period of 3 months from October 1st – December 31st 2010. The raw data was uploaded into Mysql database for the extraction of necessary features from the raw data. The features that were selected were based on those used by [12], [13] and the RFM (Recency, Frequency, Monetary) related features. These features were chosen due to the nature of pre-paid service providers. The focus is on constructing features that are able to reflect the changes in usage behavior. The final dataset used consisted of 996 subscriber call records and consisted of the following data variables which were selected from the call records in order to utilize them in building the required and targeted features:

a. Phone No of each subscriber
b. Incoming Calls
c. Incoming Start Time
d. Incoming Duration
e. Outgoing Calls
f. Outgoing Start Time
g. Outgoing Duration

2.2 Data Mining

Figure 1 presents the Data Mining framework developed for this work. Both descriptive and predictive data mining techniques were used. In the descriptive step, the customers were clustered based on their usage behavioural (RFM) feature. K-means and (EM) Expected Maximization clustering methods were used for the clustering.
K-means clustering is a partitioning method that treats observations in data as objects having locations and distances from each other. It partitions the objects into K mutually exclusive clusters, such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. Each cluster is characterized by its centroid, or centre point. EM (Expectation Maximization) assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters. EM can decide how many clusters to create by cross validation, or you may specify apriori how many clusters to generate.

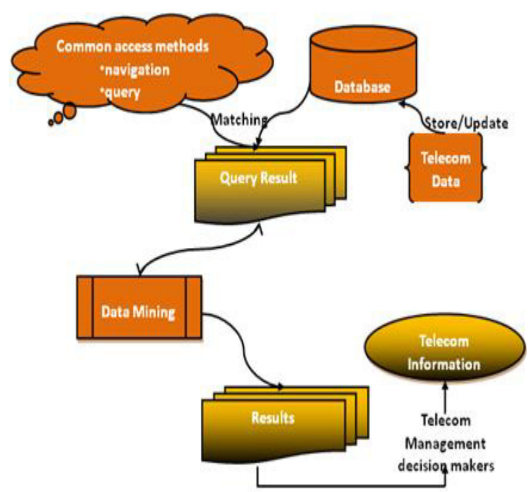

Fig 1. Data Mining Framework

For the predictive step, classification techniques were utilized. Classification is the process of finding a model (or function) that describes and distinguishes data classes or concepts for the purpose of being able to use the model to predict the class of objects whose class label is unknown. Decision tree was chosen because it is capable of efficiently generating interpretable knowledge in an understandable form. Models from tree classifier (DecisionStump, M5P, and RepTree) were used.

DecisionStump is a model consisting of a one-level decision tree. That is, it is a decision tree with one internal node (the root) which is immediately connected to the terminal nodes. A decision stump makes a prediction based on the value of just a single input feature. Sometimes they are also called 1-rules. The algorithm builds simple binary decision 'stumps' (1 level decision trees) for both numeric and nominal classification problems. It copes with mission values by extending a third branch from the stump or treating 'missing' as a separate attribute value. DecisionStump is usually used in conjunction with a boosting algorithm such as LogitBoost. It does regression (based on mean-squared error) or classification (based on entropy).

M5P implements base routines for generating M5 Model trees and rules. A learning technique that consistently yields the best results is M5P regression trees. RepTree is a fast decision tree learner. It builds a decision/regression tree using information gain/variance and prunes it using reduced-error pruning (with backfitting). The algorithm only sorts values for numeric attributes once. Missing values are dealt with by splitting the corresponding instances into pieces.

2.3 Building the Clustering Model

Weka was leveraged with Java in building the clustering model. SimpleKmeans wrapped up with MakeDensityBased Clusterer and EM (Expectation Maximization) was used for the clustering. The following set of 12 RFM related variables were constructed by the use of Mysql in order to segment the subscriber based on their calling behavior:

1. Call Ratio: Proportion of calls which has been made by each subscriber to his/her total number of calls (incoming and outgoing calls).
2. Max Date: The last date in our observed period in which a subscriber has made a call.
3. Min Date: The first date in our observed period in which a subscriber has made a call.
4. Average Call Distance: The average time distance between one's calls.
5. Life: The period of time in our observed time span in which each subscriber has been active.
6. Max-Distance: The maximum time distance between two calls of a specific subscriber in our observed period.
7. No-of-days: Number of days in which a specific subscriber has made or received a call.
8. Total-no-in: The total number of incoming calls for each subscriber in our observed period.
9. Total-no-out: The total number of outgoing calls for each subscriber in our observed period.
10. Total Cost: The total money that each subscriber has been charged for using the services in the specific time period under study.
11. Total-duration-in: The total duration of incoming calls (in Sec) for a specific subscriber in our observed time span.
12. Total-duration-outgoing: The total duration of outgoing calls (in Sec) for a specific subscriber in our observed time span.

2.4 Building the Predictive Model

Among the call details maintained in the investigated company, three measures commonly used to describe call patterns of a subscriber by aggregating his/her recall records which are:

a. Minutes of use (MOU): this refers to the total number of minutes of outgoing calls made be the subscriber over a specific period.
b. Frequency of use (FOU): this refers to the total number of outgoing calls made by the subscriber over a specific period.
c. Sphere of influence (SOI): this refers to the total number of distinctive receivers contacted by the subscriber over a specific period.

For every single cluster the following features were extracted:

a. $MOU_{initial}$: this represents the MOU of a subscriber in the first sub-period.
b. $FOU_{initial}$: this represents the FOU of a subscriber in the first sub-period.
c. $SOI_{initial}$: this represents the SOI of a subscriber in the first sub-period.
d. ΔMOU_s: this represents the change in MOU of a subscriber between the sub-period s – 1 and s (for s=2, n) and is measured by $\Delta MOU_s = (MOU_s - MOU_{s-1}+\delta)/(MOU_{s-1}+ \delta)$, where MOU_1 = MOUinitial and δ is a small positive real number (e.g. 0.01) to avoid the case when MOU_{s-1} is 0 (i.e. when ΔMOU_s cannot be calculated).
e. ΔFOU_s: this represents the change in FOU of a subscriber between the sub-period s – 1 and s (for s=2,...n) and is calculated as $\Delta FOU_s = (FOU_s - FOU_{s-1}+\delta)/(FOU_{s-1}+ \delta)$.
f. ΔSOI_s: this represents the change in SOI of a subscriber between the sub-period s – 1 and s (for s=2,...n) and is calculated as $\Delta SOI_s = (SOI_s - SOI_{s-1}+\delta)/(SOI_{s-1}+ \delta)$.

Using Decision Tree algorithms the predictive models were constructed for each of the clusters.

3. RESULTS AND DISCUSSION

The following presents the results of the descriptive and predictive models. The descriptive model was used to describe the calling behaviour of the subscribers while the predictive model was used for prediction of subscribers who are likely to churn.

3.1 Clustering (Descriptive) Model Result

In the descriptive model, SimpleKMeans and EM (Expected Maximization) algorithms were used in describing the customer behaviour. EM performed better than the SimpleKMeans algorithm. The performance measure used was the log likelihood. The log likelihood measures how well an algorithm has performed; the closer to zero the better the performance of the algorithm. From the Log likelihood value of both the SimpleKMeans clustering and EM clustering:

a. The Log likelihood value of table the SimpleKMeans is **-58.56228**
b. The Log likelihood value of table the EM is **-58.0476**

The EM result is better than that of the SimpleKMeans. Hence EM algorithm was used for the clustering model. The analysis of each of the attributes used is presented in a graphical form.

3.1.1 Call Ratio

Call Ratio is the proportion of calls which has been made by each subscriber to his/her total number of calls (incoming and outgoing calls). Almost all the clusters are having the same call ratio except for cluster 10 with 0.3. The more a subscriber calls the more likelihood he /she is retained and the more turnover for the telecom service provider. Hence the service provider should intensify effort in deploying strategy that will encourage subscribers in cluster 10 to make more calls.

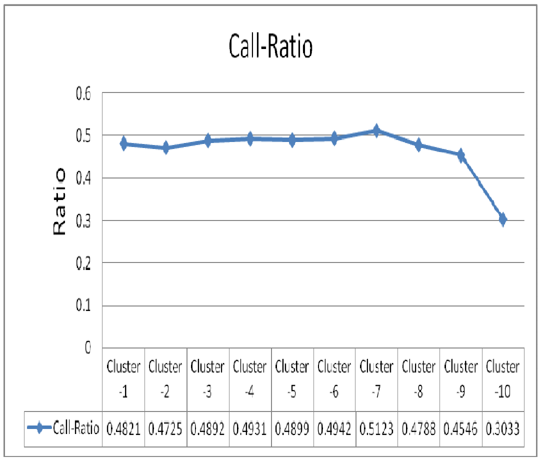

Fig 2. Call-Ratio for each Cluster using EM (Expected Maximization)

3.1.2 Average Call Distance

The Average Call Distance is the average time distance between one's calls. From figure 3, a cluster with high average call distance implies that the subscribers are not making call regularly. Clusters 8 and 9 fell into this category. This might be due to a number of reasons including getting the same service at lower cost from other service provider, quality of service to mention a few. Hence they are likely to churn in the nearest future. The telecommunications service provider should intensify retention efforts on them so as to win them back.

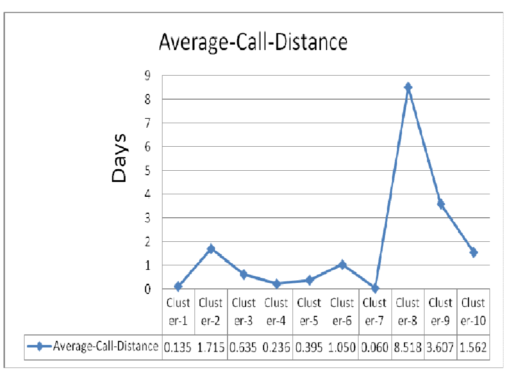

Fig 3. Average-Call Distance for each Cluster using EM (Expected Maximization)

3.1.3 Life

Life represents the period of time in our observed time span in which each subscriber has been active. Those subscribers that fall in the category of clusters 8, 9 and 10 are likely to churn which may be due to some factors such as quality of the service, coverage, price, etc. For instance if a subscriber relocates to a location where there is no network coverage, the subscriber will need to go for another network.

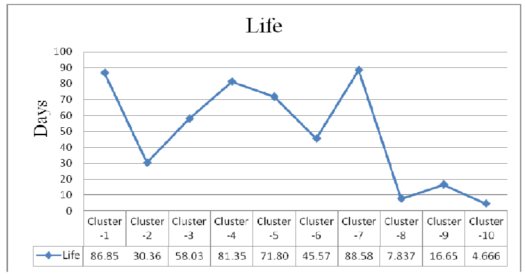

Fig 4. Life for each Cluster using EM (Expected Maximization)

3.1.4 Max Distance

Max-Distance is the maximum time distance between two calls of a specific subscriber in our observed period. Clusters with high maximum time distance represents the subscribers that have not been calling regularly. The higher the maximum time distance the more tendencies for the subscribers to churn. As a result retention efforts should be focused on the subscribers that fall into clusters 8, 9 and 10.

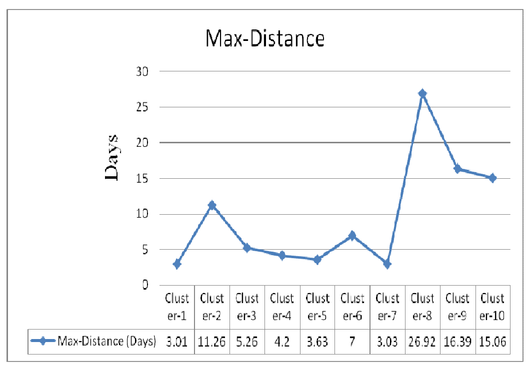

Fig 5. Max-Distance for each Cluster using EM (Expected Maximization)

3.1.5 No of Days

No-of-days stands for the number of days in which a specific subscriber has made or received a call. The total number of days in the observed period was ninety (90) days. The number of days for clusters 8, 9 and 10 is far below average and this implies that they have not been active. For them to be won back retention efforts have to be focused on them otherwise in the nearest future they are likely to churn.

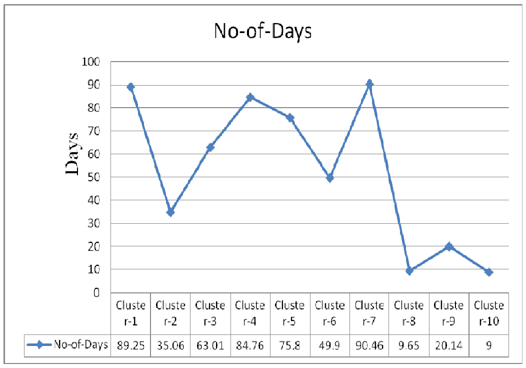

Fig 6. No-of-Days for each Cluster using EM (Expected Maximization)

3.1.6 Total No In

Total-no-in is the total number of incoming calls for each subscriber in our observed period. When a subscriber stops receiving calls though a network; it points to the fact that the subscriber might not be interested in the network again because if he/she is making calls through that network he will definitely be communicated back through that same network. Subscribers in clusters 2, 6, and especially clusters 8, 9 and 10 should be tracked to know what actually went wrong. From the investigation telecommunications service provider will now be informed of the kind of retention effort to deploy.

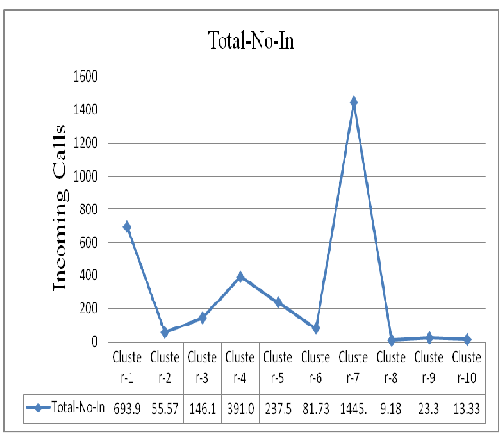

Fig 7. Total-No-In for each Cluster using EM (Expected Maximization)

3.1.7 Total No Out

Total-no-out represents the total number of outgoing calls for each subscriber in our observed period. A subscriber that stops making calls through a network will definitely not be receiving call through that network. The subscribers that are in clusters 2, 3, 6, 8-10 are likely to churn.

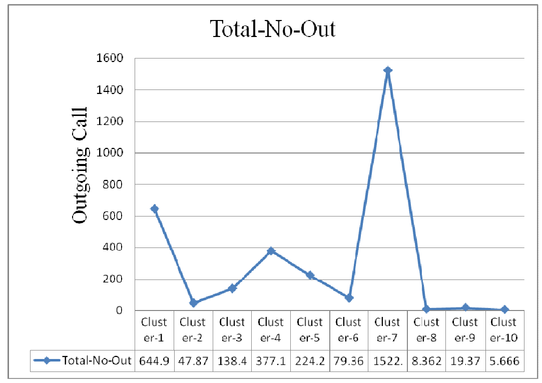

Fig 8. Total-No-Out for each Cluster using EM (Expected Maximization)

3.1.8 Total Cost

Total Cost is the total money that each subscriber has been charged for using the services in the specific time period under study. The more money spent the likelihood that the subscriber is satisfied with the network services and vice versa. Price is the most determinant factor here because the lower the price the more the total turnover for the service provider and the more calls made by the subscribers. From figure 4.8, the clusters 2, 3, 6, 8, 9 and 10 have the lowest total cost. Retention efforts should be focused on those subscribers that form those clusters.

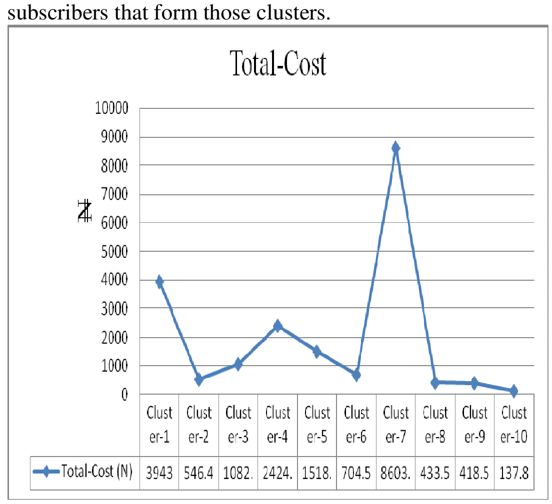

Fig 9. Total-Cost for each Cluster using EM (Expected Maximization)

3.1.9 Remarks

1. Subscribers in clusters 8, 9 and 10 have not been responding well and are very likely to churn in the nearest future. Hence they should form the focus of targeted campaign in order to win them back.
2. Subscribers in clusters 2, 3 and 6 are slightly different, their call ratio and no of days attributes were still fairly okay. With demographic data further investigation could be carried out. For instance they could be students in school. Therefore a package could be developed for them in other to encourage them.

3.2 The Classification (Predictive) Model Result

The DecisionStump, M5P, and RepTree classifier algorithm implemented in WEKA were used. Five performance measures were used in determining the performance of these algorithms on the dataset. These were:

i. Correlation coefficient (CC): This measures the degree of correlation or relationship among the attributes. It ranges between 1 for high positive correlation to -1 for high negative correlation, with 0 indicating a purely random relationship.

ii. Mean Absolute Error (MAE): This is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. As the name suggests, the mean absolute error is an average of the absolute errors. MAE can range from 0 to ∞. It is a negatively-oriented score: Lower values are better.

iii. Root Mean Squared Error (RMSE): The RMSE is a quadratic scoring rule which measures the average magnitude of the error. In other words, it represents the difference between forecast and corresponding observed values are each squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable. It can range from 0 to ∞. It is a negatively-oriented score: Lower values are better.

iv. Relative Absolute Error (RAE): This takes the total absolute error and normalizes it by dividing by average of the actual values. Lower values are better.

v. Root Relative Squared Error (RRSE): Instead of total absolute error as in RAE, it takes total squared error and divides by the average of the actual values. Finally, the square root of the result is taken.

Cluster's 8, 9 and 10 which contains the subscribers that are likely to churn now becomes the focus of the next stage where the actual churners are determined. Table 1 presents the classification algorithm results.

Table 1: Classification Algorithm Result

M5P performed better than both DecisionStump and Reptree. Hence M5P algorithm was used in building the predictive model on each cluster. The most significant features in building the predictive model for clusters 8, 9 and 10 are presented in a Table 2.

Table 2: Determinant Features for Clusters 8, 9 and 10

Cluster Number	Determinant Features
8	ΔFOUs, ΔSOIs, MOU_Final, FOU_Initial, FOU_Final and SOU_Initial
9	ΔSOIs, ΔMOUs, FOU_Initial, FOU_Final
10	ΔMOUs, ΔSOIs, MOU_Final, FOU_Initial and SOU_Initial

3.2.1 M5P Result on Cluster 8

The M5 pruned model tree (using smoothed linear models) generated just one rule which classified all 44 subscribers in cluster 8 as churner.

3.2.3 M5P Result on Cluster 9

The M5 pruned model tree (using smoothed linear models) generated the following decision tree:

```
ISOIs <= -0.527 :
|  ISOIs <= -0.619 :
|  |  ISOIs <= -0.742 : LM1 (14/23.19%)
|  |  ISOIs >  -0.742 : LM2 (55/21.703%)
|  ISOIs >  -0.619 :
|  |  IMOUs <= -0.09 : LM3 (29/32.134%)
|  |  IMOUs >  -0.09 : LM4 (10/30.915%)

ISOIs >  -0.527 :
|  FOU_Initial <= 26 :
|  |  FOU_Final <= 9.5 :
|  |  |  FOU_Initial <= 5 : LM5 (2/52.559%)
|  |  |  FOU_Initial >  5 : LM6 (7/25.64%)
|  |  FOU_Final >  9.5 : LM7 (29/24.454%)
|  FOU_Initial >  26 :
|  |  FOU_Final <= 24.5 : LM8 (13/2.656%)
|  |  FOU_Final >  24.5 :
|  |  |  FOU_Initial <= 37.5 : LM9 (14/2.494%)
|  |  |  FOU_Initial >  37.5 :
|  |  |  |  FOU_Initial <= 77 : LM10 (18/22.782%)
|  |  |  |  FOU_Initial >  77 : LM11 (5/22.962%)
```

Cluster 9 had a total of 196 instances. A total of 108 subscribers that fall under the rules LM1:LM4 were classified as churners while the remaining 88 subscribers under the rules LM5:LM11 are classified as non-churners.

3.2.5 M5P Result on Cluster 10

The M5 pruned model tree (using smoothed linear models) generated the following decision tree:

ISOIs <= -0.48 : LM1 (75/15.868%)

ISOIs > -0.48 : LM2 (20/47.175%)

A total of 75 instances were in cluster 10. Seventy (75) subscribers in cluster 10 that fall under the rule LM1 were classified as churners while the remaining 20 subscribers under the rules LM2 were classified as non-churners.

4. CONCLUSION

Inability to distinguish the churner from non-churner has been the problem of telecommunications service provider. There are two alternatives; either to send incentives to all customers (both churners and non-churners), which will be tantamount to a waste of money or to focus on acquisition program (that is, acquisition of new customers) which is more costly than retention effort. Since both alternatives have their negative implication on the finance of the company, distinguishing churner from non-churner is the best approach. Ability of the telecom service to distinguish between churner and non-churner is the central idea and the achievement of this research.

Table 3: Factor for Differentiations

	Decision Stump			M5P			RepTree		
Performance Measure	Cluster 8	Cluster 9	Cluster 10	Cluster 8	Cluster 9	Cluster 10	Cluster 8	Cluster 9	Cluster 10
CC	0.5292	0.5536	0.8739	0.9463	0.9283	0.9645	0	0.7278	0.608
MAE	0.1103	0.1896	0.1702	0.0476	0.0811	0.0866	0.1241	0.1417	0.1478
RMSE	0.1609	0.2632	0.2351	0.0613	0.1261	0.1288	0.1897	0.2171	0.386
RAE	88.84%	83.10%	77.30%	38.33%	35.57%	39.33%	99.96%	62.14%	67.13%
RRSE	84.85%	83.28%	48.61%	32.32%	39.91%	26.64%	100.06%	68.69%	79.81%

This work has been able to identify the subscribers that are likely to churn in the nearest future in one of the Nigerian Telecommunications Service providers. Specifically, the churn probability of subscribers in clusters 8, 9 and 10 is very high and hence serious retention campaign should commence otherwise those subscribers will be lost to other telecommunications service provider. The work was further able to identify the specific churners from clusters 8, 9 and 10. In-order to improve the interpretation of the results demographic data could also be added in further research works.

REFERENCES

[1] H. Kim and C. Yoon, Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. Telecommunications Policy vol. 28, 2004, pp. 751-765.

[2] R. Groth, Data mining: building competitive advantage. Santa Clara, CA: Prentice Hall, 1999.

[3] SAS Institute, Best Practice in Churn Prediction. A SAS Institute White Paper, 2000.

[4] P. Kotler and L. Keller, Marketing management. 12th ed. New Jersey: Pearson Prentice Hall, 2006.

[5] A. E. Eiben, T.J. Euverman, E. Kowalczyk and F. Slisser, Modeling customer retention with statistical techniques, rough data models, and genetics programming. Fuzzy sets, rough sets and decision making processes. Eds. A. Skowron & S. K. Pal. Berlin: Springer, 1998.

[6] C. Berne, J. M. Mugica and M. J. Yague, The effect of variety seeking on customer retention in services. Journal of Retailing and Consumer Services, vol. 8, 2001, pp. 335-345.

[7] J. Burez and D. Van Den Poel, Separating financial from commercial customer churn: a modeling step towards resolving the conflict between the sales and credit department. Expert Systems with Applications. Retrieved Nov. 15th, 2010, from http://www.feb.ugent.be/nl/Ondz/wp/Papers/wp_11_717.pdf, 2008.

[8] A. Eshghi, D. Haughton and H. Topi, Determinants of customer loyalty in the wireless telecommunications industry. Telecommunications Policy, vol. 31, 2007, pp. 93-106,

[9] S. Lee, and K. Siau, A review of data mining techniques. Industrial Management and Data Systems, vol. 101(1), 2001, pp. 41-46.

[10] A. Berson, S. Smith, and K. Thearling, Customer retention. Building data mining applications for CRM. New York: McGraw-Hill, Chapter 12, 2000.

[11] D. Hand, H. Mannila, and P. Smyth, Principles of data mining, Cambridge: MIT Press, MA, 2001.

[12] C. Wei and I. Chiu, Turning telecommunication call details to churn prediction: a data mining approach. Expert Systems with Applications vol. 23, 2002, pp. 103-112.

[13] T. J. Ali, Predicting customer churn in telecommunication service providers. Retrieved Nov. 20th, 2010, from http://LTU-PB-EX-09052-se.pdf (application/pdf object), 2009.