Essam Shaaban, Yehia Helmy, Ayman Khedr, Mona Nasr / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 4, June-July 2012, pp.693-697

A Proposed Churn Prediction Model

Essam Shaaban*, Yehia Helmy, Ayman Khedr, Mona Nasr**
*(Department of Information Systems, MUST University, 6 October, Cairo, Egypt
** (Department of Information Systems, Helwan University, Cairo, Egypt

ABSTRACT

Churn prediction aims to detect customers intended to leave a service provider. Retaining one customer costs an organization from 5 to 10 times than gaining a new one. Predictive models can provide correct identification of possible churners in the near future in order to provide a retention solution. This paper presents a new prediction model based on Data Mining (DM) techniques. The proposed model is composed of six steps which are; identify problem domain, data selection, investigate data set, classification, clustering and knowledge usage. A data set with 23 attributes and 5000 instances is used. 4000 instances used for training the model and 1000 instances used as a testing set. The predicted churners are clustered into 3 categories in case of using in a retention strategy. The data mining techniques used in this paper are Decision Tree, Support Vector Machine and Neural Network throughout an open source software name WEKA.

Keywords:-Churn prediction, classification, clustering, data mining, prediction model

1. INTRODUCTION

Churn prediction process is a highly debated research area for more than ten years. Researchers from different disciplines have tried to analyze this problem from their own perspectives to figure out a clear understanding and to recommend an effective solution for churners in many business areas. Abbasimehr et al. [1] state that churn prediction is a useful tool to predict customer at churn risk. Conventional churn prediction techniques have the advantage of being simple and robust with respect to defects in the input data, they possess serious limitations to the interpretation of reasons for churn. Therefore, measuring the effectiveness of a prediction model depends also on how well the results can be interpreted for inferring the possible reasons of churn [2]. The purpose of prediction is to anticipate the value that a random variable will assume in the future or to estimate the likelihood of future events [3]. Most DM techniques derive their predictions from the value of a set of variables associated with the entities in a database. DM models may be employed to predict customer churn developed in many disciplines such as demographic data and/or behavioral data. There are many DM techniques that can be used in classification and clustering customer data to predict churners in the near future. These techniques may use Decision Tree (DT), Support Vector Machine (SVM) in addition to Neural Networks (NN), Genetic Algorithms (GA) or Fuzzy Logic (FL) to predict churners.

This paper is organized as follows. Section 2 describes the types of churners. Section 3 shows the existing prediction models rather than the techniques of developing a predictive model. Section 4 describes the proposed churn prediction model besides the results of an implemented case study. Finally; conclusion and future work are presented.

2. TYPES OF CHURNERS

As figure 1 depicts; There are two main categories of churners which are voluntary and involuntary [4]. Involuntary churners are the easiest to identify. These are the customers that Telco decides to remove from subscribers list. Therefore’ this category includes people that are churned for fraud, non-payment and customers who don’t use the phone. Voluntary churner is more difficult to determine; it occurs when a customer makes a decision to terminate his/her service with the provider. When people think about Telco churn it is usually the voluntary kind that comes to mind.

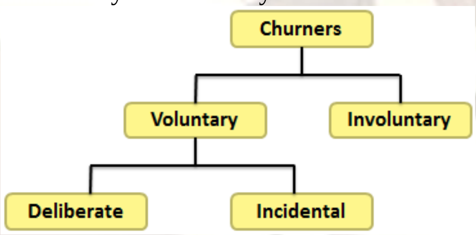

Figure 1: churn taxonomy

Voluntary churn can be sub-divided into two main categories, incidental churn and deliberate churn. Incidental churn occurs, not because the customers planned on it but because something happened in their lives. For example: change in financial condition churn, change in location churn, etc. Deliberate churn happens for reasons of technology (customers wanting newer or better technology), economics (price sensitivity), service quality factors, social or psychological factors, and convenience reasons. Deliberate churn is the problem that most churn management solutions try to solve [4] [5].

3. PREDICTIVE MODELS

Predictive modeling is mainly concerned with predicting how the customer will behave in the future by analyzing their past behavior. Predicting customers who are likely to churn is one example of the predictive modeling [6]. Predictive modeling is used in analyzing Customer Relationship Management (CRM) data and DM to produce customer-level models that describe the likelihood that a customer will take a particular action. The actions are usually sales, marketing and customer retention related. There are many models that can used to define distinguish between churners and non-churners in an organization. These models can be classified into traditional models or techniques (RA and DT) and soft computing techniques (FL and NN) [2].

3.1 Traditional techniques

3.1.1 Decision trees

DT is most popular type of predictive model. It has become an important knowledge structure, used for the classification of future events [7]. DT usually consists of two main steps, tree building and tree pruning. The tree-building step consists of recursively partitioning the training sets according to the values of the attributes. The partitioning process continues until all, or most of the records in each of the partitions contain identical values. Some branches may be removed because it could consist of noisy data. The pruning step involves selecting and removing the branches containing the largest estimated error rate. Tree pruning is known to enhance the predictive accuracy of the decision tree, while reducing the complexity [8].

3.1.2 Regression Analysis

RA is another popular technique used to deal with predicting customer satisfaction it is based on supervised learning models. Regression models deal with a dataset consisting of past observations, for which both the value of the explanatory attributes and the value of the continuous numerical target variable are known [3].

3.2 Soft computing techniques

3.2.1 Neural Networks

NN has been successfully used to estimate intricate non-linear functions. A NN is an analogous data processing structure that possesses the ability to learn. The concept is loosely based on a biological brain and has successfully been applied to many types of problems, such as classification, control, and prediction [9]. NN is different from DT and other classification techniques because they can provide a prediction with its likelihood. Various neural network approaches have emerged over time, each with varying advantages and disadvantages (Liao et al., 2004), however greater detail into these variances is beyond the scope of this paper. Research suggests that neural networks outperform decision trees and regression models for churn prediction [8].

3.2.2 Fuzzy Logic (FL)

FL is a conceptually easy to understand. The mathematical concepts behind fuzzy reasoning are very simple. Naturalness of the approach makes it preferable to the other techniques. FL is flexible, tolerant of imprecise data, and it can model nonlinear functions of arbitrary complexity. It can be blended with conventional control techniques. In many cases fuzzy systems expends the concept of the conventional control techniques and simplify their implementation. Regarding the telecom industry; there is no work achieved related to churn prediction using the fuzzy techniques [10].

There are a lot of studies have implemented in the area of telecom churn prediction. Summary of latest churn prediction studies is shown in table 1.

Table 1: churn prediction studies

Year	Author	Technique
2001	Datta et al. [11]	DT
2002	Ping and Tang [12]	DT induction
2003	Au et al. [13]	GA
2006	Ahn et al. [14]	partial defection
2007	Junxiang [15]	Survival Analysis Modeling
2008	Piotr [16]	rough-sets
2008	Seo, and Ranganathan [17]	Two-level model
2009	Jahromi et al [18]	NN and DT
2010	Gotovac [4]	DT
2011	Lee et al. [19]	partial least squares (PLS) model
2011	Yeshwanth et al [20]	Hybrid Learning
2011	Fasanghari and Keramati [21]	Local Linear Model Tree (LOLIMOT) algorithm.

4. THE PROPOSED MODEL

The proposed model is composed of six steps. As shown in figure 1, these steps are: identify problem domain, data selection, investigate data set, classification, clustering and knowledge usage. As figure 2 depicts; the classification step produces two types of customers (churners and non-churners) while the clustering step produces 3 clusters which are used to be evaluated according to the retention strategy in further usage.

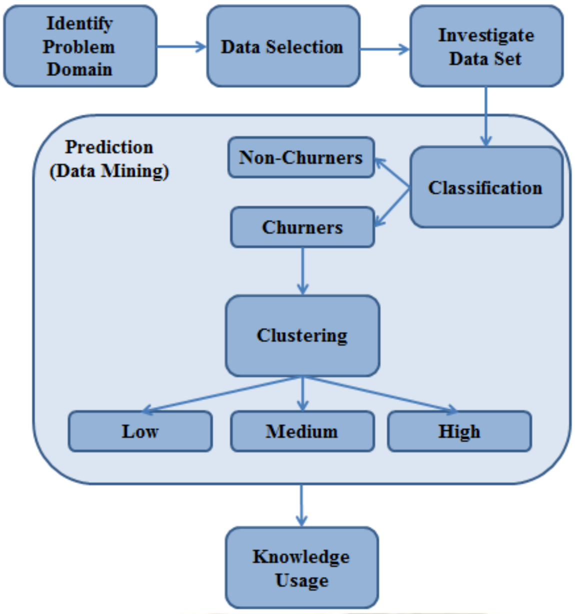

Figure 2: The Proposed Churn Prediction model

The proposed model can produce more than 3 clusters based on the types of acquired knowledge. Knowledge usage receives the produced clusters for assign a retaining solution for each type of churners. Churners can be clustered according to many criteria such as profitability or dissatisfactory of customers.

4.1 Churn Prediction with the Proposed Model

During this case study an open source DM tool named WEKA [22] is used in addition to a data set of 5000 instances. The data set is obtained from an anonymous mobile service provider. As figure 2 depicts; the data set is divided into a training set and a testing set. The training set is 4000 (80%) instances and the testing set is 1000 (20%) instances. The training data contains 3200 (80%) instances are labeled non-churners bot the others are 800 (20%) are labeled churners.

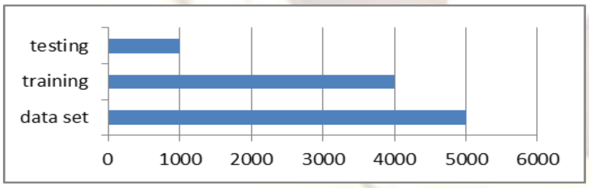

Figure 2: the distribution of the data set

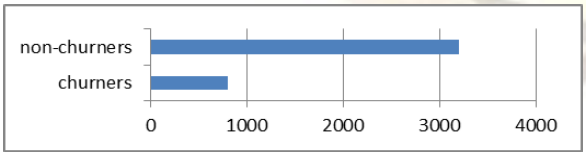

Figure 3: Training set

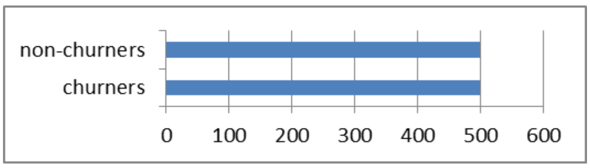

Figure 4: Testing set

The attributes for this data set are 23 attributes as shown in table 2. The class of evaluating the status of each instance in the a data set is named churn. Churn attribute is labeled churn=True if a customer left the service provider but it is labeled churn=False if he/she is still continuing with the service provider. churn=True if a customer left the service provider but it is labeled churn=False if he/she is still continuing with the service provider.

Table 2: attributes of the data set.

Attribute	Data Type	Attribute Description
Age	Number	Customer's age categorized into five groups
Gender	Number	Customer's gender (1=male, 0=female)
mar_st	Number	Customer's marital status (1=yes, 0=no)
M_in_sm_MOU	Number	mean of in minutes of use from the same service provider
M_out_sm_MOU	Number	mean of out minutes of use from the same service provider
M_sm_MOU	Number	mean of all minutes of use from the same service provider
M_in_oth_MOU	Number	mean of in minutes of use from other service provider
M_out_oth_MOU	Number	mean of out minutes of use from other service provider
M_oth_MOU	Number	mean of all minutes of use from other service provider
M_in_MOU	Number	mean of in minutes of use either from the same or other service provider
M_out_MOU	Number	mean of out minutes of use either from the same or other service provider
MOU	Number	mean of all in and our minutes of use either from the same or other service provider
chng_MOU	Nominal	Change in minutes of use from one month to another during the time of the experiment (decreased, Normal or increased)
M_sms	Number	The mean number of messages during the time of the experiment
M_M_rev	Number	Mean monthly revenues
Ass_prod	Number	Associated product (yes or no)
Ass_ser	Number	Associated services (yes or no)
M_CC_calls	Number	Mean number of customer care calls
M_drop_calls	Nominal	Mean number of dropped calls
Compaints	Nominal	Mean number of complaints
no_chng_tif_plan	Number	Number of changes in tariff plan

max_call _distance	Number	The maximum time distance between 2 calls
CHURN	Nominal	Churn=True or False

4.2 Interpretation of Results

The result of the classification step is represented in table 3.

Table 3: Confusion Matrix of results

	Actual Class	**Actual Prediction**	
		Non-churners	**Churners**
Decision Tree	Non-churners	376/500 (75.2%)	124/500 (14.8%)
	Churners	97/500 (19.4%)	403/500 (80.6%)
Neural Networks	Non-churners	420/500 (84%)	80/500 (16%)
	Churners	83/500 (16.6%)	417/500 (83.4%)
Support Vector Machine	Non-churners	416/500 (83.2%)	84/500 (16.8)
	Churners	79/500 (15.8%)	421/500 (84.2%)

The accuracy and error rate of the predicted results are shown is table 4. The accuracy rate and error rate are computed as shown in the following equations
Accuracy = (no. of correct / Total no. of predictions)
Error rate = (no. of wrong / Total no. of predictions)

Table 4: Accuracy and error rates comparison for DT, NN, SVM techniques.

Technique	**Accuracy**	**Error rate**
Decision Tree (DT)	77.9% (779/1000)	22.1% (221/1000)
Neural Networks (NN)	83.7% (837/1000)	16.3% (163/1000)
Support Vector Machine (SVM)	83.7% (837/1000)	16.3% (163/1000)

The best classification results can be extracted from tables 3 and 4 which is found in SVM classification technique that is the most relevant with NN classification technique for the data set in hand. SVM classification technique predicts 421 churners from 500. The predicted churners are used to be clustered in the next step.
During the clustering step the 421 predicted churners are used to be clustered using simple K Means algorithm. 22 attributes are used in clustering as shown in table 5. The clustering process use 3 types of clusters which one can name Cluster 0 (Low), Cluster 1 (Medium) and Cluster 2 (High) according to the suitable situation of clustering. The 3 resulting clusters can be assigned for profitability, priority for retaining, or dissatisfactory.

Table 5: Clustering Output

	3 Clusters with 421 instances		
	Cluster 0 (88) 20.9 %	Cluster 1 (208) 49.4%	Cluster 2 (125) 29.7%
Age	12	43	25
Gender	Female	Male	Female
mar_st	Single	Married	Single
M_in_sm_MOU	47	29	23
M_out_sm_MOU	12	45	12
M_sm_MOU	698	43	52
M_in_oth_MOU	41	45	65
M_out_oth_MOU	14	25	28
M_oth_MOU	77	52	158
M_in_MOU	50	54	69
M_out_MOU	730	142	85
MOU	194	232	210
chng_MOU	Decreased	Decreased	Decreased
M_sms	5	25	1
M_M_rev	21.4	112.14	29
Ass_prod	NO	NO	NO
Ass_ser	YES	NO	NO
M_CC_calls	2	6	5
M_drop_calls	0	0	0
Compaints	2	1	0
no_chng_tif_plan	0	2	3
max_call_distance	0	0	0

The output of the clustering step can be used for the knowledge usage step in order to assign a retention strategy for a specific cluster or customer. In case of clustering for a specific purpose such as customer's profitability, dissatisfactory, or cross selling this requires a specific attribute selection from table 5.

5. CONCLUSION

Many churn prediction models and techniques have been presented to date. However, a simple model is required to distinguish churners from non- churners then clustering the resulted churners for providing retention solutions. In this paper, a simple model based on DM techniques was introduced to help a CRM department to keep track its customers and their behavior against churn. A data set of 5000 instances with 23 attributes is used to train and test the model. Using 3 different techniques which are DT, SVM, and NN for classification and simple K Means techniques for clustering results indicate that the best output for the data set in hand is SVM technique. The next stage of the authors' research will involve performing a deeper analysis into the customer data to try to establish new churn prediction retention model that will use the predicted and clustered data to assign a suitable retention strategies for each churner type.

6. REFERENCES

[1] H. Abbasimehr, M. Setak, M. Tarokh. A Neuro-Fuzzy Classifier for Customer Churn Prediction.

International Journal of Computer Applications, vol. 19, no. 8, pp. 35-41, April, 2011

[2] V. Lazarov and M. Capota. Churn Prediction. Business Analytics Course. TUM Computer Science, December 2007. http://home.in.tum.de/~lazarov/files/research/papers/churn-prediction.pdf

[3] Carlo Vercellis, Business Intelligence: Data Mining and Optimization for Decision Making, John Wiley & Sons, Ltd. 2009 ISBN: 978-0-470-51138-1

[4] S. Gotovac. "Modeling Data Mining Applications for Prediction of Prepaid Churn in Telecommunication Services," vol. 51, no. 3, pp. 275-283, 2010

[5] H. Kim, and C. Yoon, "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market." Telecommunications Policy. Vol. 28 No.: PP. 751-765, 2004.

[6] M. Hassouna, Agent Based Modelling and Simulation: An Examination of Customer Retention in the UK Mobile Market. PhD thesis, Brunel University, UK, 2012.

[7] K. Muata, and O. Bryson, Evaluation of Decision Trees: AMulti Criteria Approach, Computers and Operational Research, 31, 1933-1945, 2004

[8] W. Au, C. Chan, and X. Yao, A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction, IEEE transactions on evolutionary computation, 7, 6, 532-545, 2003.

[9] R. Behara, W. Fisher, and J. Lemmink, Modelling and Evaluating Service Quality Measurement Using Neural Networks, International journal of operations and production management, 22, 10, 1162-1185, 2002.

[10] Ö. SELVİ, Traffic Accident Predictions Based On Fuzzy Logic Approach For Safer Urban Environments, Case Study: Izmir Metropolitan Area, PhD thesis, 2009.

[11] P Datta, B. Masand, D. Mani, and B. Li, Automated Cellular Modeling and Prediction on a Large Scale, Issues on the application of data mining, 14, 485- 502, 2001

[12] C. Wei and I. Chiu, "Turning telecommunications call details to churn prediction: a data mining approach," Expert Systems with Applications, Elsevier , vol. 23, no. 2, pp. 103-112, Aug. 2002.

[13] W. Au, , C. Chan, and X. Yao, A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction, IEEE transactions on evolutionary computation, 7, 6, 532-545, 2003

[14] H. Ahn, , P. Hana, and S. Lee, Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. Telecommunications Policy. Vol. 30 No.: PP. 552-568, 2006)

[15] L. Junxiang. Predicting Customer Churn in the Telecommunications Industry – An Application of Survival Analysis Modeling Using SAS. Data mining techniques. SUGI 27. Paper 114, 2007

[16] S. Piotr. Global Perspectives Mobile Operator Customer Classification in Churn Analysis. Technical University of Szczecin, Poland SAS Global Forum, pp. 1-5, 2008.

[17] B. Seo, C. Ranganathan, and Y. Babad,. "Two-level model of customer retention in the US mobile telecommunications service market." Telecommunications Policy. Vol. 32 No.: PP. 182-196, 2008

[18] A. tamaddoni, M. Moeini, I. Akbari, A. Akbarzadeh, "A dual-step multi-algorithm approach for churn prediction in Pre-paid telecommunications service providers", the 6th International Conference on Innovation & Management, SÃO PAULO, Brazil, 2009.

[19] H. Lee, Y. Lee, H. Cho, K. Im, and Y. S. Kim, "Mining churning behaviors and developing retention strategies based on a partial least squares (PLS) model," Decision Support Systems, July 2011.

[20] V. Yeshwanth, V. Vemal Raj, and M. Sharavanan, "Evolutionary Churn Prediction in Mobile Networks Using Hybrid Learning," Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference, pp. 471-476, 2011

[21] M. Fasanghari and A. Keramati. Customer Churn Prediction Using Local Linear Model Tree for Iranian Telecommunication Companies. pp. 25-37, July 2011.

[22] http://www.cs.waikato.ac.nz/ml/weka/