IEEE Compter Society Nigeria Technical Paper Series
in Collaboration with
The 26th iSTEAMS Bespoke Multidisciplinary Conference &
The Council for Scientific & Industrial Research, Ghana
The School of IT & Computing, American University of Nigeria, Yola
www.isteams.net/yola2020

DOI: dx.doi.org/10.22624/iSTEAMS/V26P1-IEEE-NG-TS
CrossREF Member Listing - https://www.crossref.org/06members/50go-

A Comparative Study of Predicting Customer Churn and Lifetime

O. Adesua
Creative Research Networks
Accra Ghama Affiliates
adesuaomolola@yahoo.com

P.A. Danquah
Council for Sci & Ind. Research
Accra, Ghana
danquahpaul@gmail.com:

O.B. Longe
American University of Nigeria
Yola, Nigeria
olumide.longe@aun.edu.ng

Abstract: ***The problem to be investigated in this research is that of predicting customers who are at risk of leaving the company, a term called churn prediction in telecommunication. The aim of this research is to predict customer churn and further focus on creating customer lifetime profiles. These profiles will allow the company to fit their customer base into n categories and make a long estimation on when a customer is potentially going to terminate their service with the company. The research then proceeds to provide a comparative analysis of neural networks and survival analysis in their capabilities of predicting customer churn and lifetime.***

.

Key words: ***GSM networks, Base station, Mobile station, Signal strength, GSM service provider***

I. INTRODUCTION

Churn' is a word derived from change and turn [42]. It means the discontinuation of a contract. The biggest revenue leakages in the telecom industry are increasing customers churn behavior. Such customers create an undesired and unnecessary financial burden on the company.This financial burden results in to huge loss of the company and ultimately may lead to sickness of the company, detecting such customers well in advance is an objective of this research.

Major concern in customer relationship management in telecommunications companies is the ease with which customers can move to a competitor, a process called “churning”. Churning is a costly process for the company, as it is much cheaper to retain a customer than to acquire a new one [43]. It is stated that the costs of acquiring a new customer is five to ten times greater than that of retaining an existing one [1]. Also, the results of Dawkins and Reichheld’s research shows that only five percent increase in retention rate yields in about 25 to 95 percent increase in the net present value of customers across a wide range of industries, such as credit cards, auto services, insurance brokerage, etc [37].

As a result of these and similar facts, enterprises are getting more and more interested in customer retention instead of acquiring new customers. Indeed, firms are concluding that the best core marketing strategy for the future is to retain existing customers and avoid customer churn ([35] and [36]).

This is exactly what churn management does. “Churn management consists of developing techniques that enable firms to keep their profitable customers and it aims at increasing customer loyalty.” [34] That is, treating customers such that “they remain client of their original supplier even if a competitor proposes more advantageous conditions.” [34]. Churn prediction models are developed by academics and practitioners to effectively manage and control customer churn in order to retain existing customers. As churn management is an important activity for companies to retain loyal customers, the ability to correctly predict customer churn is necessary.

Statistical and data mining techniques have been utilized to construct the churn prediction models. The data mining techniques can be used to discover interesting patterns or relationships in the data, and predict or classify the behavior by fitting a

model based on available data. In other words, it is an interdisciplinary area with a general objective of predicting outcomes and employing sophisticated data processing algorithms to discover mainly hidden patterns, associations, anomalies, and/or structure from extensive data stored in data warehouses or other information repositories [40] . There are several data mining techniques that are proposed to predict potential customers that are most likely to churn. Amongthe popular techniques to predict customer churn are: neural networks, support vector machines and logistic regression models [41].

The goal of most churn prediction model is to distinguish churners from non-churners as much aspossible. Although in general this approach gives satisfying results, the time aspect as to when churning will take place often is neglected. Similarly, companies are increasingly viewing customers in terms of their lifetime value.

Customer lifetime profiling therefore helps to establish customer segementation model which defines customer types in terms of lifetime value (LTV) or loyalty index. Not only do the telecommunications companies distinguish between which customers stay longer and which ones stay shorter, they also distinguish between which customers are highly profitable and which ones are low profitable or not profitable.

This research investigates a comparative analysyis of neural network and survival analysis in predicting customer churn and lifetime.

II. LITERATURE REVIEW

Building an effective customer churn prediction model using various techniques has become a decisive topic for business and academics in recent years [1]. Table 1 list relevant papers and the data mining model/technique applied in them.

Table 1: RELATED WORKS

Data Mining Technique	**References**
Neural Network	[4], [5], [6], [7], [12], [13], [14], [16], **[17]**, [22], [23], [25], [26], [27], [28]
Decision Tree	[3], [4], [6], [9], [10], **[11]**, [12], [13], [14], [16], [19], [25], [28]
Logistic Regression	[2], **[6]**, [11], [13], [18], [19], [20], [23], [28], [29], [30], [31], [33]
Random Forests	[2], **[16]**, [18], [20], [23], [31], [32]
Support Vector Machine	[2], [13], [14], [16], [18], **[19]**, [20]
Survival Analysis	**[22]**, [24], [32]
Bayesian Network	[4], [13], **[15]**
Self Organizing Maps	**[3]**, [5]
AdaCost	**[28]**
Gradient Boosting Machine	[20]
Linear Discriminant Analysis	[14]
AdaBoost	[14]
Rough Set Theory	[8]
K-Nearest Neighbor	[6]
K Means	[12]
Time series	[33]
Tailor-Butina	[33]
ROCK (Robust Clustering using linKs)	[25]
Regression Forests	[31]
Linear Regression	[31]
Association Rules	[5]
Sequence Discovery	[21]

As it is seen in table 1, a list of all reviewed references that have exploited each method is identified. Classes are sorted based on two factors: number of papers in that field and the recency of publication. For each of the top-nine techniques, an example paper, which is indicated by bolded font in the table, is briefly explained. The rationale for selecting explained papers is based on recency and benchmarking results; that is, the latest published paper in each category, in which the related model has exhibited the best performance amongst other benchmarks, is summarized.

Reference [17] uses Genetic Algorithm to improve Neural Network by optimizing the network weights. This research is conducted in the context of a mobile telecommunications company and has developed two models with fitness functions based on cross entropy (log maximum likelihood) and model accuracy. For optimizing these two models, authors have exploited the crossvalidation method.

In [11], first input variables, which are selected according to receiver operating curve (ROC) of a single input predictor, are divided into several groups (e.g. demographic, bill information, etc.) based on the concepts they describe. Then each group is used as the input of a dependent Alternative Decision Trees (ADTrees) model, which is an improvement to decision trees. The outputs of ADTrees models constitute the inputs of a Logit model that will predict the churn. Recursive feature elimination (RFE) method is exploited for eliminating ineffective features in this stage. This model is benchmarked with TreeNet model (i.e. winner of Gold Prize) on a telecom company dataset and has recorded a very similar performance according to ROC.

Reference [6] assumes a complex relationship between the independent and target variables. It proposes using K-Nearest Neighbor (KNN) method as a solution for dealing with such challenge. It first exploits a single input KNN for changing the values of each independent feature. Then the new dataset is used for training and using Logit model. The proposed method has outperformed LR, C4.5, and RBF regarding a test on four distinct datasets and accuracy and ROC criteria.

By combining the cost-sensitive learning method of Weighted Random Forests with the over-sampling technique used in Balanced Random Forests, [16] has implemented Improved Balanced Random Forests. In the new approach, repeated sampling and higher penalties for wrong classifications of members of minority class helps in selecting the best subset of input features. The chief aim of this method is to deal with class imbalance in applications like churn management. The results of comparing the new method with Weighted Random Forests, Balanced Random Forests, Decision Tree, Neural Network and SVM illustrates its better accuracy, lift and lift curve.

In [19] a two step framework for using SVM in churn prediction is suggested. In the first step, the best C (i.e. a SVM model parameter) and subset of features are chosen by applying RFE and L2-SVM techniques on the primary C candidates and set of features. Also sampling is conducted in this step by Support Vector Sampling. Then the selected variable C, the subset of features, and the primary sample are used to train a nonlinear model called RBF-SVM in the next step. Also other features of RBF-SVM are determined in this stage. The final model is the best one among the two linear and nonlinear generated models. Authors has taken the help of auxiliary methods like 10-fold cross-validation, line and pattern search in order to improve their results which have outperformed that of Logit, C4.5, L1-SVM and CART regarding AUC measure.

Reference [22] has exploited the Neural Network method to predict the hazard and survival rate of customers. For doing so, first the hazard rate is calculated by Kaplan-Meier equation. Then the resulting values are set as the target value in training a Neural Network model. The new model is named BP ANN & Survival Analysis. The second contribution of this paper is establishing a customer segmentation model which defines six customer types based on Life Time Value (LTV) and Customer Survival Phase. The latter measure is computed regarding a proposed curve, which is derived from Product Life Cycle. The authors suggest appropriate retention strategies for each segment.

In [15] the simple Bayesian method is improved by applying Genetic Algorithm (GA) to it. In fact, GA is employed to optimize the structure of the network and the weights of its edges. Comparing this method with Bayesian Networks, TAN and APRI in the context of a credit company and regarding True Positives, True Negatives, True Positive Rate and True Negative Rate, the new method shows an outstanding performance.

Reference [3] constructs a two step model in which first the probability of attrition is calculated for a customer and then, if required, an appropriate retention strategy is proposed according to the cluster that includes the customer. The authors use C5.0 and Growing Hierarchical Self Organizing Map (GHSOM) respectively in the two sequential steps. They also amend the GHSOM method and also improve C5.0 results by exploiting cross-validation based on entropy and gain rate.

Reference [28] explains a research conducted in the context of financial service. It provides a new definition for churn, which is slumping LTV. It also defines a new loss function based on the profit and uses it to propose a new model evaluation criterion called area under profit curve (AUPROC). Another contribution of this research is improving the traditional AdaBoost (i.e. a model improvement method) technique to a newer one called AdaCost, which incorporates the new loss function in calculations. Taking the AUPROC, AUC, cumulative profit percentage and model accuracy into account for different models with time horizons of 3 and 6 months, AdaCost outperforms other models.

Reference [38] identifies the variables that affect churn in reverence of customer complaints data and provides a comparative analysis of neural networks, regression trees and regression in their capabilities of predicting customer churn. It showed significant accuracy for predicting customer churn using repairs and complaints data, proving that repairs and complaints influence customer's decisions to stay with their service providers.

Reference [39], proposed a neural network (NN) based approach to predict customer churn in subscription of cellular wireless services. The results of experiments indicate that neural network based approach can predict customer churn with accuracy more than 92%.

Further, it was observed that medium sized NNs perform best for the customer churn prediction when different neural network's top ologies were experimented.

III. RESEARCH OBJECTIVES

The objectives of this reserach are to develop models to predict customer churn/survival using neural network and survival analysis. The second one is to demonstrate/test the 2 models. The final objective is compare the results from the 2 techniques.

IV. RESEARCH METHODOLOGY

Records from a Telecommunications company will be used in this research. These pertinent records will be generated from its database like billing and customer service database. Required input fields or variables will be generated from these records which will be divided into the training and testing sets. This research will be done in two stages The first stage is to generate a prediction model for each of the chosen techniques (Survival analysis and Neural Network). Each of the models will constructed using the same training dataset. Testing dataset will then be used to evaluate the perfomance of these models. The goal of this stage is to determine which customers would churn or otherwise. The second stage will involve analysing and predicting customer lifetime using the same techniques and dataset. The goal at this stage is to determine the length of time that the customers will stay with the company. The performance of the techniques are then compared in terms of accuracy in their predictions.

REFERENCES

[1] KhakAbi, S., Gholamian, M.R. and Namvar, M. "Data mining applications in customer churn management," International conference on Intelligent Systems Modelling and Simulations, 2010, pp. 2592–2602.

[2] K. Coussement and Dirk Van den Poel, "Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques," Expert Systems with Applications, vol. 34, 2008, pp. 313–327.

[3] Bong-Horng Chu, Ming-Shian Tsai and Cheng-Seen Ho, "Toward a
hybrid data mining model for customer retention," Knowledge-Based Systems, vol. 20, 2007, pp. 703–718.

[4] X. Hu, "A Data Mining Approach for Retailing Bank Customer Attrition Analysis," Applied Intelligence, vol. 22, 2005, pp. 47–60, Springer.

[5] H. S. Song, J. K. Kim, Y. B. Cho and S. H. Kim, "A Personalized Defection Detection and Prevention Procedure based on the Self-Organizing Map and Association Rule Mining: Applied to Online Game Site," Artificial Intelligence Review, vol. 21, 2004, pp. 161–184.

[6] Y. M. Zhang, J. Y. Qi, H. Y. Shu, and J. T. Cao, "A Hybrid KNN-LR Classifier and its Application in Customer Churn Prediction," Proc. the IEEE International Conference on Systems, Man and Cybernetics, Oct. 2007, pp. 3265–3269.

[7] G. Song, D. Yang, L. Wu, T. Wang, Sh. Tang, "A Mixed Process Neural Network and its Application to Churn Prediction in Mobile Communications," Proc. Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06), 2006.

[8] James J.H. Liou, "A novel decision rules approach for customer relationship management of the airline market," Expert Systems with Applications, vol. 36 (3), April 2009, pp. 4374-4381.

[9] M. Zan, Z. Shan, L. Li, L. Ai-jun, "A Predictive Model of Churn in Telecommunications Based on Data Mining," Proc. IEEE International Conference on Control and Automation, IEEE Press,2007.

[10] Yi-Fan Wang, Ding-An Chiang, Mei-Hua Hsu, Cheng-Jung Lin, Ilong Lin, "A recommender system to avoid customer churn: A case study," Expert Systems with Applications, vol. 36, 2009, pp. 8071–8075.

[11] J. Qi et al., "ADTreesLogit model for customer churn prediction," Annuls of Operations Research, vol. 168, 2009, pp. 247–265, Springer.

[12] Shin-Yuan Hung, David C. Yen and Hsiu-Yu Wang, "Applying data mining to telecom churn management," Expert Systems with Applications, vol. 31, 2006, pp. 515–524.

[13] J. Zhaoa and Xing-Hua Dang, "Bank Customer Churn Prediction Based on Support Vector Machine: Taking a Commercial Bank's VIP Customer Churn as the Example," Proc. 4th International Conference on Wireless Communications, Networking and Mobile Computing, 2008 (WiCOM'08), Oct. 2008, pp. 1-4.

[14] Y. Xie and X. Li, "Churn Prediction with Linear Discriminant Boosting Algorithm," Proc. the Seventh International Conference on Machine Learning and Cybernetics, Kunming, July 2008.

[15] Hongmei Shao, Gaofeng Zheng and Fengxian An, "Construction of Bayesian Classifiers with GA for Predicting Customer Retention," Proc. Fourth International Conference on Natural Computation, IEEE Computer Society Press, 2008.

[16] Y. Xie, X. Li, E.W.T. Ngai and W. Ying, "Customer churn prediction using improved balanced random forests," Expert Systems with Applications, vol. 36, 2009, pp. 5445–5449.

[17] P. C. Pendharkar, "Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services," Expert Systems with Applications, vol. 36, 2009, pp. 6714- 6720.

[18] K. Coussement, Dirk Van den Poel, "Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers," Expert Systems with Applications, vol. 36, 2009, pp. 6127–6134.

[19] S. Lessmann and S. Voß, "A reference model for customer-centric data mining with support vector machines," European Journal of Operational Research, vol. 199 (2), Dec. 2009, pp. 520-530.

[20] J. Burez and D. Van den Poel, "Handling class imbalance in customer churn prediction," Expert Systems with Applications, vol. 36, 2009, 4626–4636.

[21] Ding-An Chiang, Yi-Fan Wang, Shao-Lun Lee and Cheng-Jung Lin, "Goal-oriented sequential pattern for network banking churn analysis," Expert Systems with Applications, vol. 25, 2003, pp. 293–302.

[22] G. Zhang, "Customer Retention Based on BP ANN and Survival Analysis," Proc. International Conference on Wireless Communications, Networking and Mobile Computing, 2007 (WiCom), Sept. 2007, pp. 3406-3411.
[23] W. Buckinx and D. Van den Poel, "Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting," European Journal of Operational Research, vol. 164, 2005, pp. 252–268.
[24] B. Larivie`re, D. Van den Poel, "Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services," Expert Systems with Applications, vol. 27, 2004, pp. 277–285.
[25] Lian Yan, Michael Fassino and Patrick Baldasare, "Predicting Customer Behavior via Calling Links," Proc. International Joint Conference on Neural Networks, Montreal, Canada, August 2005.
[26] E Xu, S. Liangshan, G. Xuedong and Z. Baofeng, "An Algorithm for Predicting Customer Churn via BP Neural Network Based on Rough Set," Proc. the 2006 IEEE Asia-Pacific Conference on Services Computing (APSCC'06).
[27] "Predicting Customer Behavior in Telecommunications," L. Yan, R. H. Wolniewicz, R. Dodier, IEEE Intelligent Systems, IEEE Computer Society.
[28] N. Glady, B. Baesens and C. Croux, "Modeling churn using customer
lifetime value," European Journal of Operational Research, vol. 197, 2009, pp. 402–411.
[29] Jae-Hyeon Ahna, Sang-Pil Hana and Yung-Seop Lee, "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry," Telecommunications Policy, vol. 30, 2006, pp. 552–568.
[30] K. Coussement and D. Van den Poel, "Integrating the voice of customers through call center emails into a decision support system for churn prediction," Information & Management, vol. 45, 2008, pp. 164–174.
[31] B. Larivie`re and D. Van den Poel, "Predicting customer retention and
profitability by using random forests and regression forests techniques," Expert Systems with Applications, vol. 29, 2005, pp. 472–484.
[32] 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, vol. 35, 2008, pp. 497–14.
[33] A. Prinzie T and D. Van den Poel, "Incorporating sequential information into traditional classification models by using an element/position-sensitive SAM," Decision Support Systems, vol. 42, 2006, pp. 508–526.
[34] M. Lejeune, "Measuring the Impact of Data Mining on Churn Management," Journal of Electronic Network Applications and Policy, vol. 11 (5), 2001, pp. 375–387.
[35] M. Kim, M. Park and D. Jeong, The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services, *Telecommunications Policy* 28 (2004), pp. 145–159.
[36] H.S. Kim and C.H. Yoon, "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market," Telecommunications Policy, vol. 28, 2004, pp. 751–765.
[37] P. M. Dawkins and F. F. Reichheld, "Customer retention as a competitive weapon," Directors & Board, Summer 1990, pp. 42–7.
[38] John Hadden, Ashutosh Twari, Rajkumar Roy and Dymtr Ruta, "Churn Prediction using Complaints Data" World Academy of Science, Engineering and Technology 19 2006.
[39] Anuj Sharma and Kumar Pangrahi, "A Neural Network based Approach to Predicting Customer Churn in Cellular Network Services." International Journal of Computer Applications (0975 – 8887) Vol 27 – No 11, 2011.
[40] Han, J., & Kamber, M. (2001). Data Mining: Concepts and Techniques. Morgan Kaufmann.
[41] Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecomm churn management. Expert Systems with Applications, 31(3), 515–524.
[42] Lazarov, V. and Capota, M. "Churn Prediction".
[43] Jadhav, R.J., and Pawar, U. T. Churn Prediction in Telecommunication using Data mining Technology, International Journal of Advanced Computer Science and Applications, Vol 2, No. 2, February 2011.