International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:14, No:11, 2020

Churn Prediction for Telecommunication Industry Using Artificial Neural Networks

Ulas Vural, M. Ergun Okay, E. Mesut Yildiz

Abstract—Telecommunication service providers demand accurate and precise prediction of customer churn probabilities to increase the effectiveness of their customer relation services. The large amount of customer data owned by the service providers is suitable for analysis by machine learning methods. In this study, expenditure data of customers are analyzed by using an artificial neural network (ANN). The ANN model is applied to the data of customers with different billing duration. The proposed model successfully predicts the churn probabilities at 83% accuracy for only three months expenditure data and the prediction accuracy increases up to 89% when the nine month data is used. The experiments also show that the accuracy of ANN model increases on an extended feature set with information of the changes on the bill amounts.

Keywords—Customer relationship management, churn prediction, telecom industry, deep learning, Artificial Neural Networks, ANN.

I. Introduction

MANAGING customer relations is a crucial and expensive task for many markets. This is much more important for the telecommunication market where the firms have heavy investment on infrastructures. The telecommunication service providers develop strategies to have more customers to use their infrastructure capacities efficiently. Some of these strategies are aimed to get new customers and the others are on preserving customer retention. Generally, to keep customers is cheaper than to gain new ones [1]. So, the customer retention is preferred by the telecommunication service providers.

Effective management of customer retention requires the knowledge of customers satisfactions. For a while, telecommunication service providers have started to use machine learning methods to predict satisfaction levels of their customers [2]. There is a vast variety of work in the literature so far [3], [4].

Relatively old studies in the literature were based on classical machine learning techniques. Farquad et. al. [5] applied an SVM based classifier to an unbalanced data of credit card customers. Brandusoiu and Toderean [6] proposed an SVM classifier for telecommunication industry. They tried different SVM kernel functions and achieve an overall accuracy of 88 %. In [7], logistic regression and decision tree based classifiers were compared and the results showed that decision tree has much more accurate results than the logistic regression has. Although, these classical methods have some success to predict churn probabilities, they require domain knowledge and complex feature analysis.

Recently, deep learning based methods gained popularity. They generally achieve better results than the classical machine learning methods when there is a big enough data and they do not require time for feature analysis. For telecom customer churn problem, deep learning based classifiers outperform random forest classifiers [8]. In [9], an ANN based classifier is applied on different datasets for domain independent customer churn prediction. They achieved 92 % accuracy on CrowdAnaytix dataset and 71 % accuracy on Cell2Cell dataset.

In this paper, we proposed an artificial neural network based churn prediction method for telecommunication industry. The proposed method is trained and evaluated on a data set of 13,145 unique subscribers. The data set only contains the expenditure data of the subscribers and our method does not requires any personal information. This is an advantage for service providers which have separate databases for billing and subscription. We also showed that the proposed method can work for different billing durations. In the experiments, we tested our model for three different billing duration from 3 months to 9 months. The model can predict the churn classes at 83 % accuracy by using 3 months expenditure data. We also extent our original data set with the billing amount change with respect to the average billing amount.

The rest of this paper is organized as follows: Section II explains the details of the data and the artificial neural network that is used. Section III describes the system validation and experiments and we conclude our paper with Section IV.

II. Model

The proposed method is explained in two subsections. Subsection II-A describes the data set which is used for the proposed churn prediction model. In the Subsection II-B, the details of the proposed neural network model are given.

A. Data Preprocessing

The data set contains subsity and 12 month billing data of 13,145 subscribers. The original data have 19 features which 2 of them are categorical and the others are continuous. In Table I, the types and the ranges of these features are given.

It is known that the relative changes on the amounts of the bills effect the customer retention. Because of this, the feature set is extended to reflect these changes. For each month, two different features are added. These features represent the difference of the bill amount from the average bill amount

U. Vural is with the Department of Computer Engineering, Istanbul Medeniyet University, Istanbul, Turkey (e-mail: ulas.vural@medeniyet.edu.tr).

M. E. Okay and E. M. Yildiz are with Intellica Business Intelligence Consultancy Inc., Istanbul, Turkey (e-mail: ergun.okay@intellica.net, enes.yildiz@intellica.net).

and the sign of the difference. The types and the ranges of extended features are presented in Table II.

B. Artificial Neural Network Model

A detailed explanation of artificial neural networks can be found in [10]. In this subsection, only the ANN parameters used in the proposed method will be given.

The network model consists of two sequential linear layers. The first layer is a linear layer and it has 19 input nodes for the original feature set and 44 inputs for the extended set. This layer has 50 output nodes for both features sets. The linear layer uses leaky ReLU (Rectified Linear Unit) as the activation function. The network model uses batch normalization and drop out with the probability of 0.4. Output linear layer contains 50 input nodes and 2 output nodes.

In the model, Adam optimizer is preferred to use with a learning rate of 0.0005.

TABLE I
DESCRIPTION OF THE FEATURES IN THE ORIGINAL DATA SET

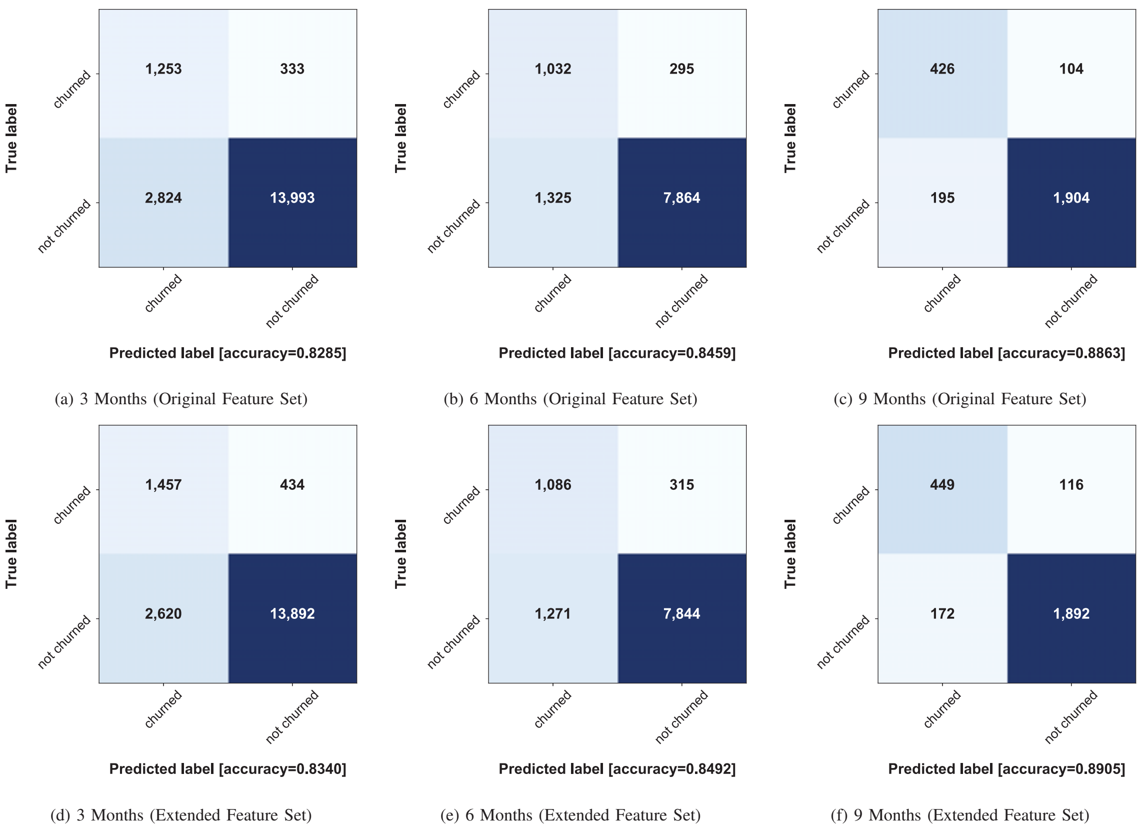
Feature	Type	Range
isPostPaid	Categorical	0, 1
isActiveInLast30Days	Categorical	0, 1
totalBillAmount (x12 months)	Continuous	[0, 14003]
numberOfDaysToActiveSubsityEnd	Continuous	[0, 725]
numberOfActiveSubsities	Continuous	[0, 1]
amountOfActiveSubsities	Continuous	[0, 4515]
numberOfSubsitiesEnded	Continuous	[0, 1]
amountOfSubsitiesEnded	Continuous	[0, 300]

TABLE II
DESCRIPTION OF THE FEATURES IN THE EXTENDED DATA SET

Feature	Type	Range
signOfBillDifference (x12 months)	Categorical	0, 1
averageBillAmount	Continuous	[0, 3907]
amountOfBillDifference (x12 months)	Continuous	[-78, 146]

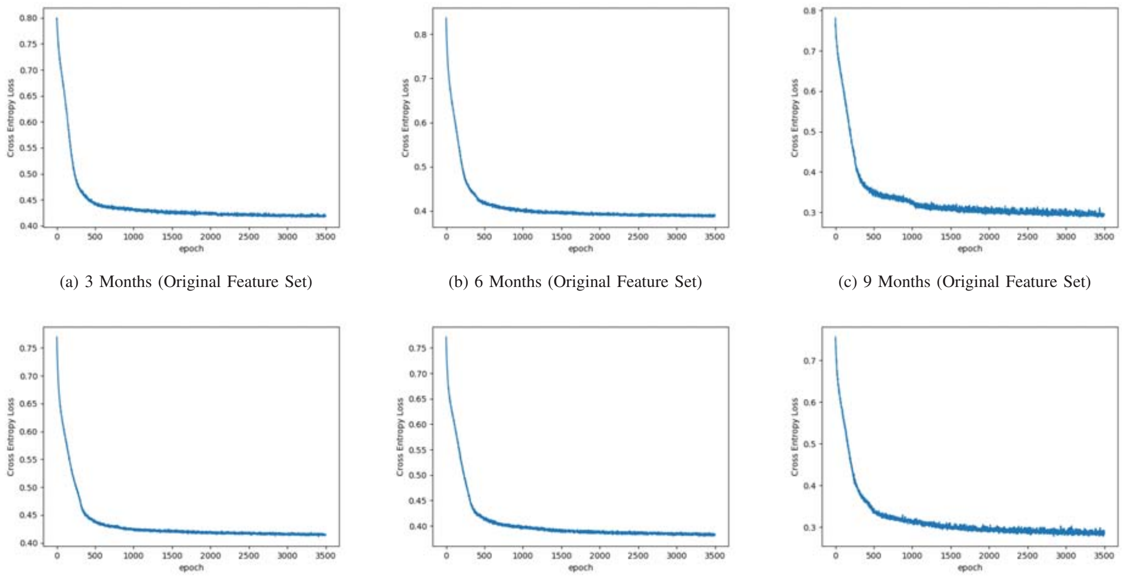
TABLE III
DISTRIBUTION OF CLASSES IN TEST AND TRAIN SPLITS

	Number Of Samples	Churners	Non-Churners
3 Months-Train	73612 (100%)	16629 (23%)	56983 (77%)
3 Months-Test	18403 (100%)	4077 (22%)	14326 (78%)
6 Months-Train	42064 (100%)	9475 (22%)	32589 (78%)
6 Months-Test	10516 (100%)	2357 (22%)	8159 (78%)
9 Months-Train	10516 (100%)	2337 (22%)	8179 (78%)
9 Months-Test	2629 (100%)	621 (24%)	2008 (76%)



(a) 3 Months (Original Feature Set)
(b) 6 Months (Original Feature Set)
(c) 9 Months (Original Feature Set)
(d) 3 Months (Extended Feature Set)
(e) 6 Months (Extended Feature Set)
(f) 9 Months (Extended Feature Set)

Fig. 1 Confusion Matrices



(a) 3 Months (Original Feature Set) (b) 6 Months (Original Feature Set) (c) 9 Months (Original Feature Set)

(d) 3 Months (Extended Feature Set) (e) 6 Months (Extended Feature Set) (f) 9 Months (Extended Feature Set)

Fig. 2 Loss values

III. Experiments

The data set used in the experiments has a billing information of 12 months but the last three months are used to determine the churn classes. The records which has no bills for the last three months are marked as churned. The data set is split into train (80 %) and test (20 %) sets. For each set, number of instances and the supports of churners and non-churners are given in Table III.

The original data set is used to create other data sets for predicting using 3 and 6 months billing data. We use all consecutive 3 and 6 months data to generate new instances.

The loss values of the training phase are shown in Fig. 2. All the models are reached to a stable loss value after 1500 - 2000 epochs. We tested our model for three different billing periods (3, 6 and 9 months) and for two different feature sets (original and extended). The results show that the model always outperforms on the extended feature set. The effect of extended features are getting more visible when the billing information shorten.

We presented several evaluation metrics to better show our results on an unbalanced dataset. Accuracy, precision, recall, F1 Score and Area Under Curve (AUC) values can be seen on Table IV. In Fig. 3, precision recall curves for original and extended feature sets are shown. As similar, ROC curves are given in Fig. 4.

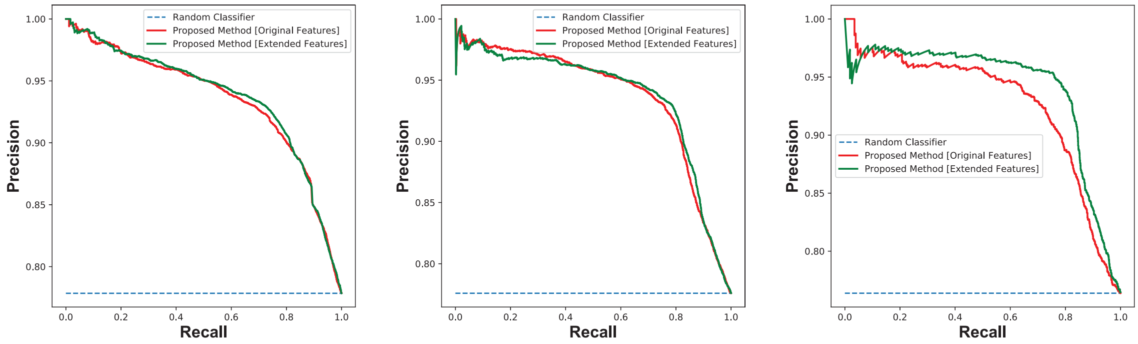
The confusion matrices are given in Fig. 1 without any normalization.

IV. Conclusion

Accurate and precise determination of customer churn has a great importance for telecom service providers. This information can be used to offer better campaigns to the customers likely to churn and it may help to increase the customer satisfaction and the brand value. The telecom companies have very large amount of customer data which allows novel artificial neural network based classifiers to work. Therefore; we can expect deep learning-based classifiers to become widespread. In this paper, we showed that billing data can be used to determine churners with only 3 month limited data. The classifier accuracy increases when it has more

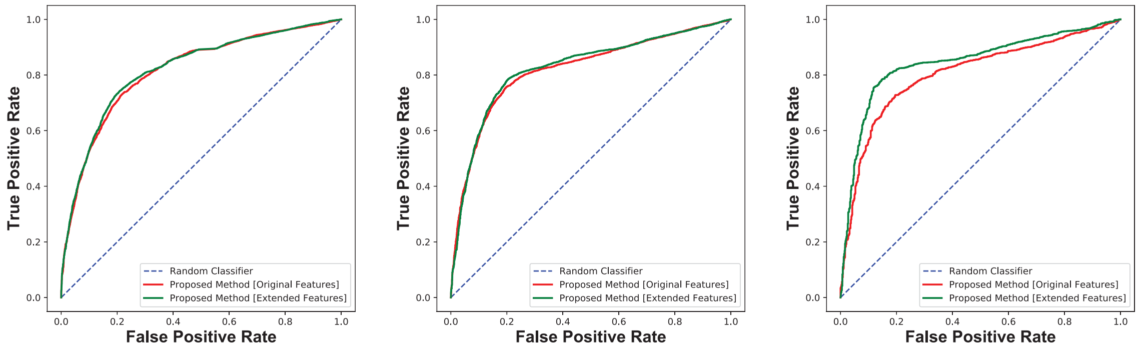
TABLE IV
Results of Evaluation Metrics

Time Window	Feature Set	Accuracy (%)	Precision	Recall	F-1 Score	AUC
3 Months	Original	82.85 %	0.22	0.86	0.895	0.937
3 Months	Extended	83.40 %	0.29	0.82	0.899	0.935
6 Months	Original	84.59 %	0.32	0.87	0.904	0.937
6 Months	Extended	84.92 %	0.37	0.83	0.906	0.937
9 Months	Original	88.63 %	0.65	0.82	0.926	0.941
9 Months	Extended	89.05 %	0.63	0.84	0.928	0.927



(a) 3 Months (b) 6 Months (c) 9 Months

Fig. 3 Precision Recall Curves



(a) 3 Months (b) 6 Months (c) 9 Months

Fig. 4 ROC Curves

information about the customers' behaviors. In the future, we plan to extend our work with a more complex ANN model which is applied to a data set with social network information.

Acknowledgment

This work is supported by Intellica Business Intelligence Consultancy Inc.

References

[1] P. Akhavan and S. Heidari, "Application of knowledge management in customer relationship management: a data mining approach," *Available at SSRN 2188230*, 2012.

[2] S. F. Sabbeh, "Machine-learning techniques for customer retention: a comparative study," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, 2018.

[3] M. Hassouna, A. Tarhini, T. Elyas, and M. S. AbouTrab, "Customer churn in mobile markets A comparison of techniques," *CoRR*, vol. abs/1607.07792, 2016. (Online). Available: http://arxiv.org/abs/1607.07792

[4] A. Ahmed and D. M. Linen, "A review and analysis of churn prediction methods for customer retention in telecom industries," in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2017, pp. 1–7.

[5] M. A. H. Farquad, V. Ravi, and S. B. Raju, "Churn prediction using comprehensible support vector machine: An analytical crm application," *Applied Soft Computing*, vol. 19, pp. 31–40, 2014.

[6] I. Brandusoiu and G. Toderean, "Churn prediction in the telecommunications sector using support vector machines," *Margin*, vol. 1, p. x1, 2013.

[7] K. Dahiya and S. Bhatia, "Customer churn analysis in telecom industry," in *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*. IEEE, 2015, pp. 1–6.

[8] F. Castanedo, G. Valverde, J. Zaratiegui, and A. Vazquez, "Using deep learning to predict customer churn in a mobile telecommunication network," *Wise Athena LLC*, 2014.

[9] V. Umayaparvathi and K. Iyakutti, "Automated feature selection and churn prediction using deep learning models," *International Research Journal of Engineering and Technology (IRJET)*, vol. 4, no. 3, pp. 1846–1854, 2017.

[10] I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, "Artificial neural networks," *Cham: Springer International Publishing*, p. 39, 2017.