Survey of Churn Customer Prediction In Business Using Neural Network

W.S.Kanmani[1], Mrs. B. Jayapradha[2]
[1, 2] Dept of Computer Science
[1, 2] Dr. Ambedkar Government Arts College, Chennai - 600 039

Abstract- *Background: In most of recent research works are evolve around with neural network. So the neural network has enormous developing in every field because Neural Network learns to do task by observational data without task specification programs. Scope and propose of neural network are increasingly popular in business sectors like marketing, forecasting, telecommunication, banking and finance. Maintenance of valuable customer are important for business, using neural network in business forecasting becoming effective. Method: Neural network has been identified as a better technology for prediction, classification, quantitative modeling, and statistics analysis in Business forecasting, due to its effective performance on identifying relationship between non-linear data. Findings: To achieve best prediction of non-profitable customer are lead to good establishment for business in future Improvements: our future work it to prediction of default customer for banking sectors*

Keywords- Churn Customer, MLP, ANN, SOM, IBRF.

I. INTRODUCTION

Over the last decade, neural networks have found application across a wide range of areas from business, commerce and industry. it solves wide range of problems better than traditional approach due do this, many of business organization are investing in neural network. The different between neural network and statistical approach is that neural network make no assumptions about the statically distribution or properties of the data, and therefore tend to be more useful in practical situations. Neural networks are also an inherently nonlinear approach giving them much accuracy when modelling complex data patterns.

There are several types of neural networks, each with a different purpose, architecture and learning algorithm. Neural networks form the backbone of most of the data mining products available, and are an integral part of the knowledge discovery process which is central to the methodology. This paper focus on how the different neural network methodology combine with data mining and prediction churn customer. According to a recent study [2], approximately 95% of reported neural network business application studies utilise multilayered feed forward neural networks (MFNNs) with the back propagation learning rule. This type of neural network is popular because of its broad applicability to many problem domains of relevance to business

Burez and Van den Poel (2007) indicate that there are two types of targeted approaches to managing customer churn: reactive and proactive. When a company adopts a reactive approach, it waits until customers ask the company to cancel their service relationship. In this situation, the company will offer the customer an incentive to stay. On the other hand, when a company adopts proactive approach, it tries to identify customers who are likely to churn before they do so. The company then provides special programs or incentives for these customers to keep the customers from churning. Targeted proactive programs have potential advantages of having lower incentive costs. However, these systems may be very wasteful if churn predictions are inaccurate, because companies are wasting incentive money on customers who will not churn. Therefore, it is important to build a customer-churn prediction model as accurately as possible

II. LITERATURE REVIEW

The authors focused on multi layer preceptron (MLP) to learn for prediction in telecommunication because churn represent the problem of losing customer to another business competitor which leads to serious lost in profit. most previous work are focus on machine learning to predict churn customer but this research Identifying important factors can greatly support customer relationship management in telecommunication companies to plan effective customer retention strategies. For that reason, in this research, we use a multilayer preceptron neural network not only for predicting customer churn but to give an insight on the relative importance of each input variable regarding customer churn.

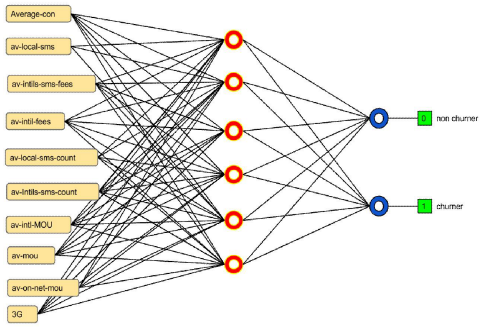

Figure 1. Multilayer Preceptron Neural Network (MLP)

The author investigates two different approaches based on neural networks for identifying important variables. The first is based on error change and the second is based on weights contribution in the network.

Limitations:

The approach is that noise and redundancy in the training set can lead to degradation in its reliability.

This author comparing two methods for weight adjustment both of them gives same amount of response.

Authors uses a feed-forward back-propagation network whose topology (number and configuration of nodes in the hidden layer) is based on the number and types of the input and output fields. To prevent over-training problems that can occur within neural networks, randomly selected proportion of the training data is used to train the network. As the data pass repeatedly through the network, it is possible for the network to learn patterns that exist in the sample only and thus over-train. So the network may become too specific to the training sample data and loose its ability to generalize.

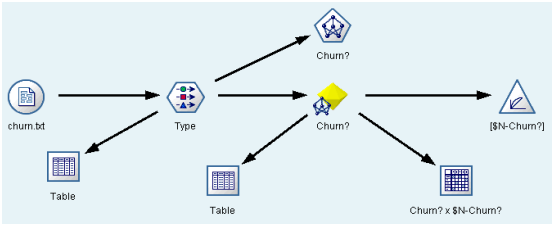

Figure 2. Prediction Accuracy of the testing and FTD

Table 1. Matrix of Actual (Rows) and Predicted (Columns) Churn

		$N-Churn?	
Churn?		False	True
	Count	0	1
	Row %	0.000	100.000
False	Count	2038	50
	Row %	97.605	2.395
True	Count	114	224
	Row %	33.728	66.272

The randomly selected proportion of the training data is used to train the network and once this proportion of data has made a complete pass through the network, the rest is reserved as a test set to evaluate the performance of the current neural network architecture.

The author considers two hybrid models by combining two different neural network techniques for churn prediction which are back-propagation artificial neural networks (ANN) and self organizing maps (SOM). The hybrid models are ANN combined with ANN (ANN + ANN) and SOM combined with ANN (SOM + ANN). In particular, the first technique of the two hybrid models performs the data reduction task by filtering out unrepresentative training data. Then, the outputs as representative data are used to create the prediction model based on the second technique.

The author creates three models for training and testing the dataset.

- Base line

At first, we use the original dataset to train a MLP neural network as the baseline ANN model for comparisons.
Table 2. The best parameter setting of the MLP baseline model.

Testing set	Learning epochs	Hidden nodes	Accuracy
Subset 1	50	32	92.80
Subset 2	300	32	92.75
Subset 3	50	8	91.81
Subset 4	100	16	88.08
Subset 5	50	12	86.99

Table 3. Prediction performance of ANN + hybrid models.

Testing set	Learning epochs	Hidden nodes	Accuracy
Subset 1	50	32	94.32
Subset 2	300	32	93.70
Subset 3	50	8	93.68
Subset 4	100	16	93.71
Subset 5	50	12	90.18

- ANN+ANN

The first hybrid model is based on cascading two ANN models, in which the first one performs the data reduction task and the second one for churn prediction.

- SOM+ANN

For the second hybrid model, a self-organizing map (SOM), which is a clustering technique, is used for the data reduction task. Then, the clustering result is used to train the second model based on ANN.

Limitation:

The data spilt into subset and it train and tested separated this is tedious process. In a baseline model the best parameter need to be set to achieve accuracy.

In hybrid models the first component (ANN, SOM) are used to filters un related this gives hectic task compare to data mining process.

Author proposed a novel learning method called improved balanced random forest (IBRF), and uses to demonstrate in churn prediction and investigate the effectiveness of standard random forest approach prediction in churn customer, also to achieve better performance than most existing system by integrating sampling techniques and cost-sensitive learning into the approach. This paper combine two methods weighted and balance random forest. The sampling technique used in weighted random forest is more efficient to handling imbalanced data and tolerance. The cost sensitive learning applied in balance random forest has more efficient on decision learning methods.

The algorithm takes as input a training set D={(X1,Y1),……,(Xn,Yn)}, where Xi,i=1,….n is a vector of descriptors and Yi is the corresponding class label. The training set is then split into two subsets D+ and D-, the first of which consists of all positive training samples and the second of all negative samples.

Advantage

The experiment result shows that the IBRF algorithm produces higher accuracy than exiting algorithms

Limitation

IBRF employs internal variables to determine the distribution of samples. Although the results are found to be insensitive to the values of these variables, imposing some limitations.

Author applies two stages of evaluation the first stage uses k-means clustering method for data reduction and form first classification. In second stage Genetic Programming (GP) applied on selected classification model to develop final classification model.

To evaluate the developed model, the author refers to the confusion matrix which is source for accuracy estimation in classification. Based confusion matrix two criteria is used first one identifies the percentage of the total number of predictions correctly classified and second is the actual churners rate percentage which identified correctly. The data set is split into three mutually exclusive subsets, and each subset is used as testing while the rest is used as training and the process is repeated three times. Eventually for both training and testing the entire data set is used.

Limitation

Limitation of this GP algorithm is accuracy rate that does not give any information about how well non-churners and churners are classified. It just measures the overall accuracy.

The result comparison of ANN, C4.5 GP algorithm is shows that the accuracy rate prediction is better result in ANN compare to all.

Table 4. Prediction results for GP and K-means+G.

Parameters	**Fold 1**		**Fold 2**		**Fold 3**	
Accuracy (%)	**Actual churners rate (%)**	**Accuracy (%)**	**Actual churners rate (%)**	**Accuracy (%)**	**Actual churners rate (%)**	**Accuracy (%)**
K-means+GP	8.67	75.80	82.30	74.60	91.40	70.20
GP	92.70	68.50	92.60	58.00	92.10	67.40
C4.5	94.10	73.60	94.10	71.70	94.10	73.60
ANN	92.20	57.30	93.50	67.40	94.80	76.90

Table 5 shows the summary of Churn Customer Prediction

S.no	Authors Name	Prediction Methods	Advantages	Limitations
1	Omar Adwan, Hossam Faris, Khalid Jaradat, Osama Harfoushi and Nazeeh Ghatasheh[5]	Multilayer Preceptron Neural Network	Two different approaches to identifying important variables. CoE and weights contribution in the network	The approach is that noise and redundancy in the training set can lead to degradation in its reliability

2	Anuj Sharma Dr.PrabinKumar Panigrahi[9]	Feed Forward Back Propagation	prediction accuracy rate is increased	The best result are obtain from one hidden layer with three neuron
3	Chih-Fong Tsai a, Yu-Hsin Lu[3]	Back propagation neural network , self organization model(hybrid NN)	hybrid models performs the data reduction task by filtering out un-representative training data	The data spilt into subset and it train and tested separated it is tedious process. best parameter setting need to be set to achieve accuracy.
4	RubaObiedat, Mouhammd Alkasassbeh, Hossam Faris and Osama Harfoushi[12]	K-means clustering , Genetic programming	Based confusion matrix two criteria is used first one identifies the percentage of the total number of predictions correctly classified and second is the actual churners rate percentage which identified correctly.	Accuracy rate that does not give any information about how well non-churners and churners are classified. It just measures the overall accuracy
5	Yaya Xie , Xiu Li , E.W.T.Ngai, Weiyun Ying[14]	Improve Balanced Random Forest	. The experiment result shows that the IBRF algorithm produces higher accuracy than exiting algorithms	IBRF employs internal variables to determine the distribution of samples. Although the results are found to be insensitive to the values of these variables, imposing some limitations

III. CONCLUSION AND FUTURE WORK

There are many challenges and neumors ways to predict the churn customer in communication industry. Most of the paper proposed using the feed forward neural network with various activation levels to achieve better results. In future we apply these concepts to predict default customer in banking sector.

REFERENCES

[1] AlisaBilalZorić Interdisciplinary Description of Complex Systems “PREDICTING CUSTOMER CHURN IN BANKINGINDUSTRY USING NEURAL NETWORKS” vol 14(2), pages116-124, 10 March 2016.

[2] Sumathi, S. and Sivanandam, S.N.: Introduction to Data Mining Principles. Studies in Computational Intelligence 29(3), 1-20 2013,

[3] Domingo, R.: Applying data mining to banking. Http://www.rtdonline.com, accessed 18th November 2015,

[4] Burez, J., & Van den Poel, D. (2007). Crm at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. Expert Systems with Applications, 32, 277–288.

[5] Omar Adwan, Hossam Faris, Khalid Jaradat, Osama Harfoushi and Nazeeh Ghatasheh. Predicting Customer Churn in Telecom Industry using MLP Neural Networks: Modeling and Analysis. Life Sci J 2014;11(3):75-81

[6] Chiang, D., Wang, Y., Lee, S., & Lin, C. (2003). Goal-oriented sequential pattern fornetwork banking churn analysis. Expert Systems with Applications, 25(3),293–302.

[7] Nazari, M. and Alidadi, M.: Measuring credit risk of bank customers using artificial neural network. Journal of Management Research 5(2), 17-27, 2013,

[8] Salvatore Madonna, Prof,Greta Cestari, PhD the accuracy of bankruptcy prediction models: a comparative analysis of multivariate discriminant models in the italian context European Scientific Journal December 2015 edition vol.11, No.34 ISSN: 1857 – 7881 (Print) e - ISSN 1857- 7431.

[9] Anuj Sharma,Dr.Prabin Kumar Panigrahi. A Neural Network Based Approach for Predicting Customer Churn in Cellular Network Services.International

Journal of Computer Applications (0975 – 8887)Volume 27– No.11, August 2011

[10] Tsai, C.F. and Wu, J.W.: Using neural network ensembles for bankruptcy prediction and credit scoring. Expert Systems with Applications 34(4), 2639-2649, 2008.

[11] Wilson, R.L. and Sharda, R.: Bankruptcy prediction using neural networks. Decision Support Systems 11(5), 545-557, 1994.

[12] Omar Adwan, Hossam Faris, Khalid Jaradat, Osama Harfoushi and Nazeeh Ghatasheh. Predicting Customer Churn in Telecom Industry using MLP Neural Networks: Modeling and Analysis. Life Sci J 2014;11(3):75-81

[13] Deep Learning. Yann LeCun, Yoshua Bengio & Geoffrey Hinton. Nature 521, 436–444 (28 May 2015)

[14] Yaya Xie , Xiu Li , E.W.T. Ngai, Weiyun Ying, Customer churn prediction using improved balanced random forests,

[15] Ivakhnenko, Alexey (1971). "Polynomial theory of complex systems". IEEE Transactions on Systems, Man and Cybernetics (4): 364–378.

[16] Mohsen Nazari , Mojtaba Alidadi “Measuring Credit Risk of Bank Customers Using Artificial Neural Network” Journal of Management Research, ISSN 1941-899X,2013, Vol. 5, No. 2.

[17] Bar, M.V.: The Computational Intelligence Techniques For Predictions-Artificial Neural Networks. Annals of Computational Economics 2(42), 184-190, 2014