

Literature Review of Using DWHBI Approaches to Predict and Reduce Customer Churn in Telecommunications Industry

K. Prakash Krishnan¹, Dr. A. Kumar Kombaiya²

¹Ph.D Scholar, Department of Computer science, Chikkanna Govt Arts College, Tirupur, Tamil Nadu, India

²Assistant Professor, Department of Computer Science, Chikkanna Govt Arts College, Tirupur, Tamil Nadu, India

ABSTRACT

Article Info

Volume 6, Issue 5

Page Number: 176-180

Publication Issue :

September-October-2020

In recent days, telecom industry plays a vital role in our daily life. During corona period entire world depends on the telecom services domain. But telecome industry has been facing many survival problems in the global market since last 10 years due to heavy competition between competitors. To stand in this field, service providers have to understand the complete customer requirements and provide the efficient services to stop the customer movement from one network to another network. Customer churn is one of the most critical problem faced by the telecom industry. In this industry, it is more expensive to bring the new customers as compared to retain the existing customers. The objective of customer churn prediction is to find the subscribers that are ready to move from the current service provider in advance. Churn prediction allows the service providers to offer new benefits and campaign offers to retain the existing customer in the same network. Technically this term would be call it as 'Win back Situation' in telecom industry. The high volume of data generated by telecom industry , with the help of data warehosuing and business intelligence implementation would become the main asset for predicting the customer churn. To prevent the churn many models and methods are used by researchers. In this paper, we reviewed different mining methods and the most popular algorithms which used in telecom industry. But this model is not only for telecom domain, it can be implement for other domains which has highly depends on customer interactions.

Keywords : Churn prediction, Churn prevention, data mining, Data warehouse and Business Intelligence, Key performance indicators [KPI model], Big data, Hive Ingestion platform, Kafka, Customer retention, campaign offers, churn model score, customer life cycle process, Performance metrics, Predictive models, ARPU

Article History

Accepted : 01 Oct 2020

Published : 07 Oct 2020

I. INTRODUCTION

India is the world's 2nd largest telecommunication market. The telecom market can be divided into 3 categories.

- 1) Wire-less
- 2) Wire-line
- 3) Internet Services

The total number of telephone subscriber in India reached 1.18 billion as of 31 March 2020. As per TRAI the number of wireless subscribers is over 1.17 billion and the number of wireline subscribers are 20.58 Million.

As of 2019, India holds the highest data usage per smartphone average 9.8 GB per month. It is expected to double by 2024 as 18 GB.

This telecom sophisticated industry is giving mammoth profits and revenues to network operators. But still it is facing lots of ups & downs to stabilize the continued growth of the product. Country like India which has the largest populations and it is not such easy to promote offers to grab the attentions.

Customer churn – It is the process of customer switching from one service provider to another service provider in given time. Retaining the existing subscribers is more profitable than fetching new subscribers. The customer churn leads to huge loss to the company and it affects company status in the global market.

ARPU [Average Revenue Per User] – A measure of the revenue per unit or user. It helps to identify the revenue generators based on the products. ARPU is a key performance indicator for service providers and used as an important benchmark strategy by investors and analyst.

Formula of ARPU = Total Revenue / Average Subscribers

ARPU is calculated either monthly (or) quarterly (or) annually time period. But it can be calculated for any interval period of time. Average subscribers can change constantly.

High ARPU will generate high revenue to the service providers. Low ARPU will generate low revenue to the company likewise high churn rate would cause the low average revenue per user.

The main objective of customer churn is used to predict the customers with high tendency to leave a company. In order to retain the existing customers , The telecom industry needs to find out the reason for churn, which can be extracted based on customer usages and customer behavioral statistics. Customer demographic attributes (age, sex, education, status, location) are used for subscriber churn calculation. It will be used to predict the prompt accuracy of newly developed model.

Literature discovers following type of customer churns.

- A) Volunteer churn - Customer quit the contract and move to the next service provider
- B) Non – Volunteer churn - Company quit the service to a customer
- C) Rotational Churn - Customer can terminate the contract without any prior knowledge of both parties (Customer and Company). Sometime it call it as Silent churn.

II. REASONS FOR CUSTOMER CHURN

The first two type [1.A] [1.B] of churn can be predicted easily with the help of customer usages and

log histories but third type [1.C] of churn is quit difficult to predict because such type of customers can terminate the contract at any time in near future. Below are some reasons for the customer churn.

- Poor customer service
- Network coverage
- In sufficient network towers
- Hidden charges and Rates
- Unwanted Spam messages
- Weak customer relationship
- Ignoring customer complaints
- Over due SLA
- Difficult user experience
- Giving much attention on Postpaid subscribers than Prepaid subscribers
- Bandwidth congestion
- Best offers compared to other networks
- Customer collaboration
- Non profitable plans
- Forced on Higher Rates
- Improper Bill cycles
- Too many advertisements
- DND feedback.

Aforementioned points are affected service provider revenue in the global market and finally it enforced the company to get the higher churn rate at the end of the month. It would lead the huge loss for the company, share holders & partners.

III. LITERATURE SURVEY

Bo Jin, Jia shi [1] proposed a new set of approach called PBCCP neural network algorithm which executes particle classification optimization (PCO) and particle fitness calculation (PFC). PCO classifies the categories according to their fitness but PFC calculates the fitness values of each forwarding training process of Back Propagation NN (Supervised Learning Algorithm). Here gradient descent is used to

minimize the error function. The actual performance of BP on a specific problem is dependent on this input data. This experiment demand us to do multiple time until we reach the desired result.

$$W_x = W_x - a \left(\frac{\partial \text{error}}{\partial W_x} \right)$$

Asif & Kadan [2] discovered the approach based on machine learning and big data platform. In order to measure the performance they used Area Under Curve (AUC) and Social Network Analysis (SNA). Here SNA has given the significant performance compared to AUC. This model was developed and tested through hadoop framework environment which handles large set of dataset. This model experimented with 4 Algorithms: Decision Tree, Random Forest, Gradient Boosted Machine (GBM), Extreme Gradient Boosting (XGBOOST). Here the data model needs to be updated each time. They concluded XGBOOST gives the best result with 93% accuracy.

Pamina J, Beschi Raja, Sathya Bama [3] divulged the method which based on machine learning process and the objective of this churn prediction is to perform binary classification with subscriber records and to figure out who are likely to cancel the contract in near future. Performance of the model evaluated by various measures like Accuracy, Recall, Precision, F-Score, ROC & AUC. This model experimented with KNN model, Decision Tree, Random Forest model, NN model. After experimental they found decision tree is the best model for real time business problem. Here they additionally PCA model for preprocessing and reduce the dimension of the data which is important from the whole dataset. Adnan Amin, Sajid Anwar [4] proposed and intelligent rule-based decision making method based on Rough set theory to remove the dispensable attribute from the given dataset. RST approach applied on different rules generation algorithms (EA, GA, CA and LA). Here they discovered genetic

algorithm dominated well than other rules generation algorithms in terms of precision, recall, the rate of misclassification, lift, coverage, accuracy, and F-measure. Below RST concepts are involved to reduce the data set

- a. Lower Approximation
- b. Upper Approximation
- c. Boundary Region
- d. Outside Region
- e. Reduction core

Maurus Riedweg , Pavol Svaba and Gwendolin Wilke [6] discovered the semi auto predictive analytics and campaign management. They have chosen large data set from switzerland telecom company and they performed the calculation based on Naïve bayes prediction model and corresponding nomogram values and this prediction accuracy exceeded decision tree approach as well as the currently used bench mark.

V.Umayaparvathi , K.Iyakutti [7] have divulged with general churn prediction models as data collection, preparation, classification and prediction. They have analyzed with existing clustering and classification methods are used to distinguishing churners and several standard performance metrics are proposed for predicting and analyzing the performance model.

- $ACC = (TP + TN) / (TP + TN + FP + FN)$
- $Precision = TP / (TP + FP)$ $Recall = TP / (TP + FN)$
- $F1 = 2 TP / (2TP + FP + FN)$

IV. CONCLUSION

In this paper , we have reviewed the detailed process of churn prediction and prevention system in telecom industry. From the literature survey conducted and clearly understood that churn

prediction in telecom industry is very much important and it would be the heart of the business in telecommunication department. Customer churn prediction and customer retention is an inevitable program in every sectors. The purpose of this paper is not to propose a new algorithm but it mainly focused on the existing algorithm and approaches. There are many different approaches to predict the customer churn process and new models are being discovered on daily basis in all other fields not only telecommunication sector. Future research direction is planned in the direction for developing a new hybrid framework to predict and reduce the customer churn process. Good processing models can be developed with mixture of existing data mining techniques and KPI values.

V. REFERENCES

- [1] Yu, R., An, X., Jin, B. et al. Particle classification optimization-based BP network for telecommunication customer churn prediction. *Neural Comput & Applic* 29, 707–720 (2018)
- [2] Ahmad, A.K., Jafar, A. & Aljoumaa, K. Customer churn prediction in telecom using machine learning in big data platform. *J Big Data* 6, 28 (2019)
- [3] Pamina J., Beschi Raja J., Sam Peter S., Soundarya S., Sathya Bama S., Sruthi M.S. (2020) Inferring Machine Learning Based Parameter Estimation for Telecom Churn Prediction. In: Smys S., Tavares J., Balas V., Ilyyasu A. (eds) *Computational Vision and Bio-Inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing*, vol 1108. Springer, Cham
- [4] Adnan Amin, Sajid Anwar (2016) Custome churn prediction in Telecom using Rough set approach *Neuro computing* 237(2017)242-254. 2016 Elsevier B.V. All rights reserved.

- [5] K.Kim , J.Lee improved churn prediction in telecom industry by analyzing a large network , Expert syst. Appl 41 (15) (2014) 6575-6584
- [6] Maurus Riedweg , Pavol Svaba and Gwendolin Wilke (2018) Revenue Optimization of Tele marketing campaigns for prepaid customers. 18714.1PFES-ES.
- [7] V.Umayaparvathi , K.Iyakutti (2016) A survey on customer churn prediction in Telecom industry : Datasets, mehtods and metrics ISSN: 2395 -0056

Cite this article as :

K. Prakash Krishnan, Dr. A. Kumar Kombaiya, "Literature Review of Using DWHBI Approaches to Predict and Reduce Customer Churn in Telecommunications Industry", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 5, pp. 176-180, September-October 2020. Available at doi : <https://doi.org/10.32628/CSEIT206535>
Journal URL : <http://ijsrcseit.com/CSEIT206535>