# CUSTOMER PREDICTION ON TELECOM USED MACHINE LEARNING IN BIG DATA PLATFORM-REVIEW

Radhika Patil[1]

[1] Assistant Professor, Department of Computer Science and Engineering, BIET, Davanagere
Email: ps.radhika@gmail.com[1]

**Abstract**: **Customers are the foundation of any business success and that is why firms are aware of itthe importance of gaining customer satisfaction. Customer churn is important issue and is considered one of the most important issues among firms because of increased conflict between firms, increased importance of marketing strategies as well customers have behaved in recent years. Organizations need to develop differently strategies for solving complex issues based on the services they provide. Customer fraud practice is essential for competition and rapid development in the telecommunications sector. The process Transfers from a service provider to another call service provider occur as a result of good services or prices or because of the various benefits offered by the competitor company customers when registering. Because of the high costs associated with acquiring new customers customer churn predictions have been developed as a very important part of planning**

**Process and strategic decision-making in the telecommunications sector. The main purpose of the research is to do check customer predictions on telecom using big machine learning data platform. Machine learning methods have been used to measure customer opportunities to meet. This study uses retrospective and KNN data for larger details to predict consumer fraud in the communications sector. Retract of goods used extensively evaluating churn opportunities as a function of set flexibility or features of customers. Similarly, with churn K-Nearest Neighbour is used to check whether a customer churns or not depending on the proximity of its feature to customers in each category. This is a lesson uses Kaggle's database website for predicting and analyzing churn. Research findings indicate that the predictive accuracy rate in the consumer churn is found to be 0.80 percent and the area below the curve is found to be 0.71 percent.**

## 1.1 Background

The global telecommunications industry is becoming one of the major sectors and consequently the growth of technology and the number of continuous operators increase the level of competition. Telecom firms are making a living competitors' market and alternatives are designed to bring in a large amount of funds. Improving customer retention time is important for companies to do reduces the likelihood of a customer conflict, referred to as "customer departure from one service provider to another service provider (Ascarza, et al 2016). The motto of customersis considered a major obstacle in the service sector with growing cutthroat services (Ahmad, et al 2019). Many studies (Umayaparvathi and Iyakutti 2016; Jutla no Sivakumar 2005) emphasized that machine learning programs are on the rise effectively predicting this situation.

The basic principle of customer prediction in the telecommunications industry is that list subscribers who probably feel they would like to leave the company used so far and I suggest solutions to prevent large churns. Recently, making measurement of churner before they stop is now required in the solid state competition between companies. The major role played by the telecommunications industry did most importantly to create predictive mechanisms along the churn lines prediction. Few studies show the ultimate user value in this industry. One of research shows that 1% reinforcement on the customer keeps movement possible in fact it increases by 5% increase in

all company shares (Kisioglu and Topcu 2010). Huang, et al (2012) predicted that in the field of telecommunications, Churn's monthly customer base is 2.2% (Yildiz and Albayrak 2018) this year. The churn customer rate was 27%. Consumer retention in the telecommunications industry has already become a nightmare as a result expanded competitive services. Brandusoiu and Toderean (2016) proposed improvement a data mining method to obtain customer churn using machine learning algorithms namely NN (Neural Networks) and SVM (Support Vector Machine). Findings emphasized that machine learning algorithm works better in customers 'predictions explosion. Yena et al. (2009) used a machine learning model to solve the churn problem in the middle customers at major telecommunications firms. Huang et al. (2015) investigated the issue of churn between customers in the area of the big data platform, which aims to emphasize that big data greatly improves the continuation of churn guessing based on variety, speed and volume of records. Ahmad, et al (2019) clarified that it is social. The features of a network analysis app improve churn in telecommunications sector. Among all the other sectors, for decades the telecommunications industry has been experiencing magnitude every year churn rate ranges from 20 to 40 percent (Ahna, et al 2006). This is important financial results for the firm, as it carries five to ten times as much as new delivery user rather than keeping an older user within the firm (Junxiang Lu, 2003). Modern companies they aim to build strong relationships with their users (Chen et al., 2012). As a result, it is it has become the belief that the best advertising policy to keep users older or older customer churn dealing (Mohammadi et al., 2013). Hashmi, et al (2013) emphasize that machine learning methods work with large, large, linear data sets with improved predictive accuracy but are thought to be more complex in respect of real world applications. Artificial Neural Networks (ANN) is a widely used algorithm method used dealing with complex problems, for example the problem of churn prediction (Mozer, et al 2000). Vector Machines (SVM) Support Vector, developed by Boser, et al (1992) supervised learning algorithm that scans data and identifies patterns, used specifically for postponement and separation. Decision Trees (DT) is another way

to use machine learning without amazing performance in non-line detection as well complex relationships between factors. Au et al (2003) emphasized that Neural networks in comparison may produce better performance than Decision Trees. However, depending on the churn problem among customers, the level of the accuracy of the Decision can be improved, depending on the data form (Hadden, et al. 2006). Similarly, the division of Naïve Bayes has shown positive results in terms of the emerging speculation in the telecommunications sector (Nath, et al. 2003) and in comparison is also possible achieve better prediction levels (Asthana, 2018) than other widely used algorithms, such as as DT. KNN and logistic regression algorithms were selected for this study because KNN the algorithm can perform segmentation work without prior knowledge about the data Distribution while order is used to estimate the potential for an explosion such as flexible set performance or customer features. Your comparison of both algorithm will help predict customer churn accurately and resolve error things that lead to customer retention. K-Nearest Neighborhood and Logistic Regression Neighborhood the algorithm exceeds other algorithms in terms of configuration data sets. Comparisons between machine algorithm models were made by Khan, et al (2010). Studies have shown that the neural network reaches more slowly than decision trees and asset losses. Vafeiadisa, et al (2015) made comparisons customer prediction analysis and findings have shown that decision trees are also with BPN reached 94% accuracy, followed by SVM (93%) while Logistic The deferral reached a low accuracy level of 86%. According to customer speculation, SVM methods have been extensively tested and estimated that they can grow predictably.

## 1.2 Comparison of various algorithms in customer churn

Yabas and Chankya (2013) explain that customer fraud has become a major issue concern for customer management most mobile service providers as the result of the costs associated with it. Current research describes our work in the customer churn investigation and evaluation of such services. The study used data mining the ability to accurately and successfully predict subscribers who will change and eventually switch to another service provider or competitor for a similar or related

service. A lesson make an effort to find alternatives or improvements high scores were recorded for efficient and effective use of resources. The the paper also focuses on the collection of separately investigated meta-classifiers and selected in accordance with their performance.

Talwar and Dahiya (2015) pointed out that customer prediction is one of the key factors features of current telecom customer relationship management systems. This text is valid with the modern system of machine learning involves deceptive predictions. This study performs frequent ML chemical testing that separates churn customer patterns in the field of telecommunications. Challenges of the sections addressed to the telecom churn predicting, the current paper makes an effort to propose a variety of hybrid methods, where classification systems are grouped into categories and performance pre-processing methods. Research shows that machine learning classes work best when available sufficient human effort made by feature engineering; as a result, it is possible to find a sensible border of classes in the feature space. Methods of machine learning it will be an effective way to make the churn process more efficient and automatic prediction.

Das and Gondkar (2018) argue that customer loyalty is considered a backlog, i.e. increasingly leading to customer frustration due to the competitive world, the flexibility of price and strengthening benefits from a competitive company. In this study, research on different algorithms of machine learning in addition to the problems of predictions of attracting customers are shown. Studies have shown that a large amount of ML The churn prediction algorithm has been seen in the telecom industry and Support Vector Machine (SVM), ANN widely used churn analysis algorithm. It's also obvious from the study that a two-step combination of the ANN training process once the comparative method of SVM testing provides a better definition level with Extended area under the Curve (AUC) than other available strategies. It's over again note that both ANN and SVM have a better chance of predicting churn and are likely be employed at the right level of prediction.

Jayaswal, et al (2016) emphasized the expansion of land and market liberalization as something that enhances market competition significantly. The the emergence of technologies present in business processes has increased competition as well set new service provider challenges. Incident of discontinuation the company by the user is called churn and, in this framework, it predicts the customer. The purpose of fraud is known as churn prediction. Both data mining and machine learning algorithms, when applied to consumer behaviour and data usage may contribute to this churn management procedures. The current paper monitors customer use and is relevant data from telecom service providers to investigate fraud in the telecommunications sector. The decision trees and their species and advanced Gradient trees are used as essential mathematical machine learning strategies for developing binary split churn. The part of the simulation is done using an apache spark composed of integrated data an analytical structure for machine learning. Achieving better and more effective outcomes, The use of a grid-focused hyper-parameter is used.

Radosavljevik, et al (2013) showed that different methods of machine learning including tree trunks, NB, NN and genetic algorithms, which are widely used algorithms for developing tabular churn predictive models. Reducing customer churn measure, it is important for telephone operators to develop self-detection and recovery systems proper encouragement for users with a growing tendency to be cunning. This study looks at a file of the extent to which social features are found in the built-in graph customer communication can be used to increase churn prediction authenticity in the pre-paid category. Research has highlighted that instead of aiming at all future churners, we can reduce our resources by focusing only on churners with the power went up.

Madan, et al (2015) aims to investigate recent publications in the field of telecommunications the customer jumps heavily with two ideas, i.e., the method used in churn telecom prediction and year of publication. Studies have emphasized neural highly profitable networks can remember the database on the basis of incomplete or noisy data. As a result, they work well with solid data operations. Research has found that rising payloads are one of the key factors an area of concern for the telecommunications sector as a whole. In such a competitive environment, companies could not

deal with the problem of non-payment. To deal with this devastation users, mining process and machine learning data can be monitored. The study does an effort to find certain interesting structures, which can be the basis for them It can detect the causes of fraud and it is possible to predict who is more it may soon become too proud.

Gupta, et al (2018) proposed a hybride-model calculator learning process distortion on mobile phone networks. Research makes an attempt to define a machine learning methods always use churn.

Lemmens and Gupta (2013) used motivational models for Stochastic Gradient in analysis churn prediction in the telecommunications industry. Kaur and Mahajan (2015) used data mining and R-tools for predicting churn rate in the telecommunications industry. Gursoy, (2010) used data mining tools to predict fraudulent customer performance in the telecommunications sector and found it successfully. The author has used data mining techniques as a decision tree and asset retreat. Adebiyi, et al (2016) used asset management in learning customer churn and a decision to be retained in the telecommunications industry. Neural networks, partition trees and the retreat was used by Poel and Lariviere (2004) in predicting what happened to the customer.

Hadden, et al, (2006) used Matlab and Dwivedi, et al (2019) hired SAS Enterprise Miner for churn prediction on telecom. Or, machine learning found is effective in customer prediction, most studies have used data mining once electronic learning methods have been overlooked. Therefore, the present study makes an effort to fill this research gap with the use of machinery used to predict customer entry telecommunications.

There are three basic research objectives. Another test "customer churn telephone prediction using a large machine learning data platform." Second, it was worth it analyze the importance of the churn prediction model in the media industry to help the telecommunications industry to prevent the rate of customers in its very first phase. Third, it was worth it compare effective algorithms for lowering the churn rate in the telecom industry. In this case, various algorithms related to the churn prediction of comparable customers and we discussed its effectiveness. The research has had a positive impact on the telecommunications industry we propose sound solutions to reduce the churn level in the telecommunications industry.

## References

[1]. Siddalingaih, Neelambike&Baraki, Parashuram. (2018). Cluster based authentication for VANET for improving the security. Compusoft. 7. 2556-2559. 10.6084/ijact.v7i2.699.

[2]. Siddalingaih, Neelambike& J, Chandrika. (2018). A DECENTRALIZED ADAPTIVE MEDIUM ACCESS CONTROL FOR V2I VANET. ICTACT Journal on Communication Technology. 9. 1696-1701. 10.21917/ijct.2018.0248.

[3]. Siddalingaih, Neelambike& Chandrika, J. (2018). An efficent distributed medium access control for V2I VANET. Indonesian Journal of Electrical Engineering and Computer Science. 9. 742-751. 10.11591/ijeecs.v9.i3.pp742-751.

[4]. Siddalingaih, Neelambike& Chandrika, J. (2017). AN EFFICIENT ENVIRONMENTAL channel modelling in 802.11p mac protocol for v2i. ICTACT journal on communication

[5]. Adebiyi, S.O., Oyatoye, E.O., Amole, B.B. (2016). Relevant Drivers for Customers` Churn and Retention Decision in the Nigerian Mobile Telecommunication Industry. Journal of Competitiveness, 8(3), pp. 52-67.

[6]. Ahmad, A.M., Jafar, A, Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(28), pp. 1-24.

[7]. Ahna, J.H., Hana, S.P., Lee, Y.S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry", Telecommunications Policy, 30(10– 11), pp. 552–568.

[8]. Almana, A.M., Aksoy, M.S., Alzahrani, R. (2014). A Survey on Data Mining Techniques in Customer Churn Analysis for Telecom Industry. Int. Journal of Engineering Research and Applications, 4(5), pp. 165-171.

[9]. Alwin, P.K.D.N.M., Kumara, B.T.G.S., Hapuarachchi, H.A.C.S. (2018). Customer Churn Analysis and Prediction in Telecommunication for Decision Making.

2018 International Conference on Business Innovation (ICOBI), pp. 40- 45.

[10]. Amin, A, Anwar, S, Adnan, A, Nawaz, M, Alawfi, K, Hussain, A, Huang, K. (2016). Customer churn prediction in the telecommunication sector using a rough set approach. Neurocomputing, volume 237, pp. 242-254.

[11]. Asthana, P. (2018). A comparison of machine learning techniques for customer churn prediction. International Journal of Pure and Applied Mathematics, 119(10), pp. 1149-1169.

[12]. Azeem, M, Usman, M. (2018). A fuzzy based churn prediction and retention model for prepaid customers in telecom industry. International Journal of Computational Intelligence Systems, 11(1), pp. 66 – 78.

[13]. Brandsoiu I, Toderean, G, Ha, B. (2016). Methods for churn prediction in the prepaid mobile telecommunications industry. In: International conference on communications, pp. 97–100. x

[14]. Chen, Z.Y., Fan, Z.P., Sun, M. (2012). "A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data", European Journal of Operational Research, volume 223, pp. 461–472.

[15]. Dahiya, K, Bhatia, S. (2015) Customer churn analysis in telecom industry. IEEE Publications.

[16]. Dong, R, Su, F, Yang, S, Cheng, X, Chen, W. (2017). Customer Churn Analysis for Telecom Operators Based on SVM. International Conference on Signal and Information Processing, Networking and Computers, volume 473, pp. 327-333.

[17]. Dwivedi, A, McGaugh, M. (2019). Telecom Industry: Customer Churn Prediction. Retrieved on 2nd December 2019 from https://www.lexjansen.com/wuss/2019/177_Final_Paper_PDF.pdf

[18]. Ebrah, K, Elnasir, S. (2019). Churn Prediction Using Machine Learning and Recommendations Plans for Telecoms. Journal of Computer and Communications, 7(11).

[19]. Gupta, M, Gandhi, A.B., Gupta, S.C. (2018). Machine Learning as Intelligent tool for Churn Prediction in Telecommunication Industry. International Journal of Computer Applications, 181(10), pp. 16-22.

[20]. Gursoy, U.T.S. (2010). Customer churn analysis in telecommunication sector. Istanbul University Journal of the School of Business Administration, 39(1), pp. 35-49.

[21]. Hadden, J., Tiwari, A., Roy, R &Ruta, D. (2006). Churn prediction: Does technology matter? World Academy of Science, Engineering and Technology, volume 16, pp. 193-199.

[22]. Hashmi, N, Butt, N.A., Iqba, M. (2013). Customer Churn Prediction in Telecommunication A Decade Review and Classification. IJCSI, volume 10, pp. 271-282.

[23]. Hassouna, M., Tarhini, A., Elyas, T., &AbouTrab, M. S. (2016). Customer churn in mobile markets a comparison of techniques. arXiv preprint arXiv:1607.07792. 20. He, Y, He, Z, Zhang, D. (2009). A study on prediction of customer churn in fixed communication network based on data mining. In: Sixth international conference on fuzzy systems and knowledge discovery, volume 1, pp. 92–94.