Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

12

Telecom Customer Churn Prediction: A Review

Srushti Lohiya, Omkar Salunkhe, Samiksha Parshionikar, Prof. Surabhi Thatte

Department of Computer Science and Applications, School of Computer Science & Engineering,

Dr. Vishwanath Karad MIT World Peace University, Pune, India.

***Abstract*—**

In recent years, predicting customer-churn in the telecom sector has been one of the most popular study subjects. It entails identifying clients who are inclined to revoke their service subscriptions. The mobile telecommunications market has undergone a transformation from one of rapid growth to one of saturation and intense rivalry. Since these customers are more likely to migrate to the competitor in the near future, telecommunications companies are now more focused on keeping their current clients. The process of creating a robust and reliable churn prediction model takes time, but it is crucial. This paper provides an excellent overview of customer churn, including its impacts, causes, consequences for businesses, methodology, and all churn prediction strategies. It comprises a wide range of methodologies proposed by previous studies as well as the technology used in these studies. New researchers will be able to find all the data they require for their churn prediction model requirements in one place thanks to this study. This report provides a thorough analysis by thoroughly outlining the research that has been done in the area and will act as a vast knowledge base for all predictions of churn in the telecom industry.

I. INTRODUCTION

Today's communication technology industry is very competitive. Customer turnover is presently a critical issue affecting essentially all telecoms' sectors globally. The telecommunications paradigm defines churn as the process by which consumers leave an

organization and stop using its services due to dissatisfaction with the services and superior offerings from other network providers within the customer's affordable price range. The corporation can suffer a loss of earnings as a result of this. Keeping customers has also grown to be a challenge. In order to build an effective churn prediction model, various elements are taken into account, including customer behavior data, the technique employed, among other things, feature selection and customer social networks. These components help create a churn-prediction model and increase its utility. After the model has been constructed, it is required to assess its performance to the performance criteria.

This paper's goal is to give researchers a platform to make this process easier and, in turn, spend less time and effort on such tasks. It also gives a very thorough introduction to churn and churn prediction, as well as the effects of churn prediction on various businesses and the causes of churn. The study covered many approaches to churn prediction in literature. Before creating a churn prediction model, it is crucial for the telecom industry to have a thorough grasp of the dataset. This report provides a comprehensive overview by listing all available datasets, together with information on their size and other properties. that have been employed by prior studies. Additionally, the report describes the various attribute types that can be found in provided telecom datasets.

The goal of the churn analysis [7] is to identify among the customers who will discontinue using a product or service. Additionally, a data mining-based project called the customer churn study will be used to uncover these possibilities. Due to today's intense competition, many businesses are now offering the same product at remarkably similar levels of service and quality.

By giving each client a likelihood, the Churn-Analysis [8makes it possible to predict events with precision in which consumers will discontinue using services or goods. According to consumer segments and the magnitude of the loss, this study can be carried out (monetary equivalent). These evaluations can be used to inform how to better communicate with customers in order to influence them and win their loyalty. The churn rate, also known as customer attrition, can be used to design marketing efforts that are effective for your target audience. Profitability can thus be greatly increased or potential damage from client loss can

be decreased at the same rate. Churn rate, for instance, is measured as 10% if a service provider with 2-million subscribers gets 750.000 new members while losing 275.000 clients. The number of customers a company loses has a big impact on how much it is worth financially. Therefore, the majority of businesses monitor their client value on a monthly or quarterly basis.

II. ANALYSIS OF CUSTOMER CHURN PREDICTION METHODOLOGIES

The bulk of the previous churn prediction systems produced incredibly accurate forecasts using algorithms using meta-heuristics and machine learning. While some authors aimed at improving data samples through efficient pre-processing and the inclusion of social characteristics [8], feature extraction, and selection techniques, others focused on using algorithms like SVM and ANN. These sorts of surveys are essential for coming up with new solutions to the age-old churn issue [9].

The need for balanced data for customer attrition prediction was emphasized by J. Burez et al. [1]. By using sampling (random and sophisticated undersampling), to improve the accuracy of churn-prediction, balancing is done between a cost-sensitive learner (weighted-random forests) and boosting (gradient boosting machine).

Yet, the usefulness of one of these methods—random variable selection—is affected. Veronika Effendy et al. [2] offered a workable solution to the problem of processing unbalanced data in order to improve customer churn forecasting. To improve the precision of churn prediction, the suggested method uses sampling and Weighted-Random Forest (WRF) to balance the dataset. The sampling procedure employs both SMOTE and undersampling. The data is classified using WRF for precise churn prediction, the fundamental step requires sampling for issues with data imbalance. The combined sampling method results in higher F-measure and accuracy values, illustrating how fewer data records are needed for accurate prediction. Despite the performance being rather respectable, the basic under-sampling technique is not very noteworthy.

To improve churn and detect insurance fraud, G. Ganesh Sundarkumar et al. [3] introduced under-sampling using a oneclass SVM. Before employing machine learning techniques for classification, the data are first undersampled using one-class SVM. The findings

demonstrate that the decision-tree surpasses competing classification methods and, when combined with one-class SVM, lowers system complexity while increasing prediction accuracy.

The present customer churn prediction methodologies, according to Qiu Yihui et al. [4], cannot fully meet the application demands because they lack a set of scientific, systemic, and procedural underpinnings. The authors suggested a feature selection method based on the orientation-ordering pruning method (OOPM). The classifier combination is pruned using this method rather than handling attribute selection. The second phase involves using higher-level consumer data to extract various features using a feature extraction method called FE RF&T. The results of the evaluation show that the FE RF&T with OOPM enhances churn prediction.

Feature-based Churn-prediction Improvement

A paradigm for complementing the fusing of multilayer characteristics to boost churn prediction rate was established by Qiuhua Shen et al. [5]. The proposed architecture mostly used feature factorization and feature generation to integrate features.

Our method improves the precision of churn forecasts by addressing the issue of large dimensionality and unbalanced data. Sebastián Maldonado et al. suggested a helpful method for choosing features based on the profit model in [6]. This strategy focuses on choosing the most important traits for the classifier stage. The feature parameters are carefully chosen for profit, much as the SVM-classifier is built on a profit foundation. The method enables more flexible usage of the kernel functions for improved prediction accuracy. Yet, SVM as a fundamental classifier does not abide by the law.

Machine-Learning Methods

To increase churn prediction's precision, SVM for structural risk minimization was suggested by Xia Guo-en et al. [7]. The suggested method concentrates on foreseeing infrastructure vulnerabilities and establishing a link between them and customer turnover. The main advantages include high churn rates, less missing records, excellent precision regardless of how many attributes, and non-linearity data. The weight of the customer sample data and the

selection of the kernel function, however, are flawed. High dimension and non-linear time sequences are also not handled correctly.

Yaya Xie-et al proved balanced-random forests (IBRF)-based churn prediction was put forth in their study by [30]. This approach incorporates sampling techniques and cost-sensitive learning while using random forests to anticipate churn. Performance limitations arise, however, because time-varying variables really aren't taken into account while making predictions.

Bayesian networks (BN), neural-networks (NN), support-vector machines (SVM), and other machine-learning methods have all been employed by Ionut Brândusoiu et al. to predict churn [11]. Using Multi-Layer-Perceptron (MLP), SVM, and BN, the authors examined data from the telecom industry. The dataset is first preprocessed using Principal Component Analysis (PCA) before machine learning categorization. According to evaluation results, SVM delivers a higher degree of accuracy compared to MLP and BN. The main reason for concern is that particular, efficient algorithms rather than ensemble approaches or mixed machine learning are utilized to predict churn.

Preeti K. Dalvi [12] suggested a churn-prediction method combining decision-trees and logistic-regression. The suggested approach is based on integrating data-mining and machine-learning strategies and evaluating their effectiveness side by side. Based on the rules and strategies, The decision tree offers a visual representation of the available facts, and logistic regression is used to measure the effect of each variable on the choice to churn. The evaluation's results demonstrate that using this method increases prediction accuracy. The technique has the disadvantage of having a very tiny class universe, however, it also cuts down on the time needed for churn prediction.

Hybrid Churn Prediction Methods

In order to predict churn in virtual settings, Hsiu-Yu Liao et al. [13] devised an approach. The authors of this paper used a hybrid-classification model that included machine learning and meta-heuristics. When determining consumer behavior, It considers and incorporates the financial cost, user behavior, and social- neighbor traits.

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

As a result, the proposed model forecasts the churn precisely and rapidly. Even though churn prediction is improved by the hybrid-model with merged features, the multi objective problem develops when various features are considered.

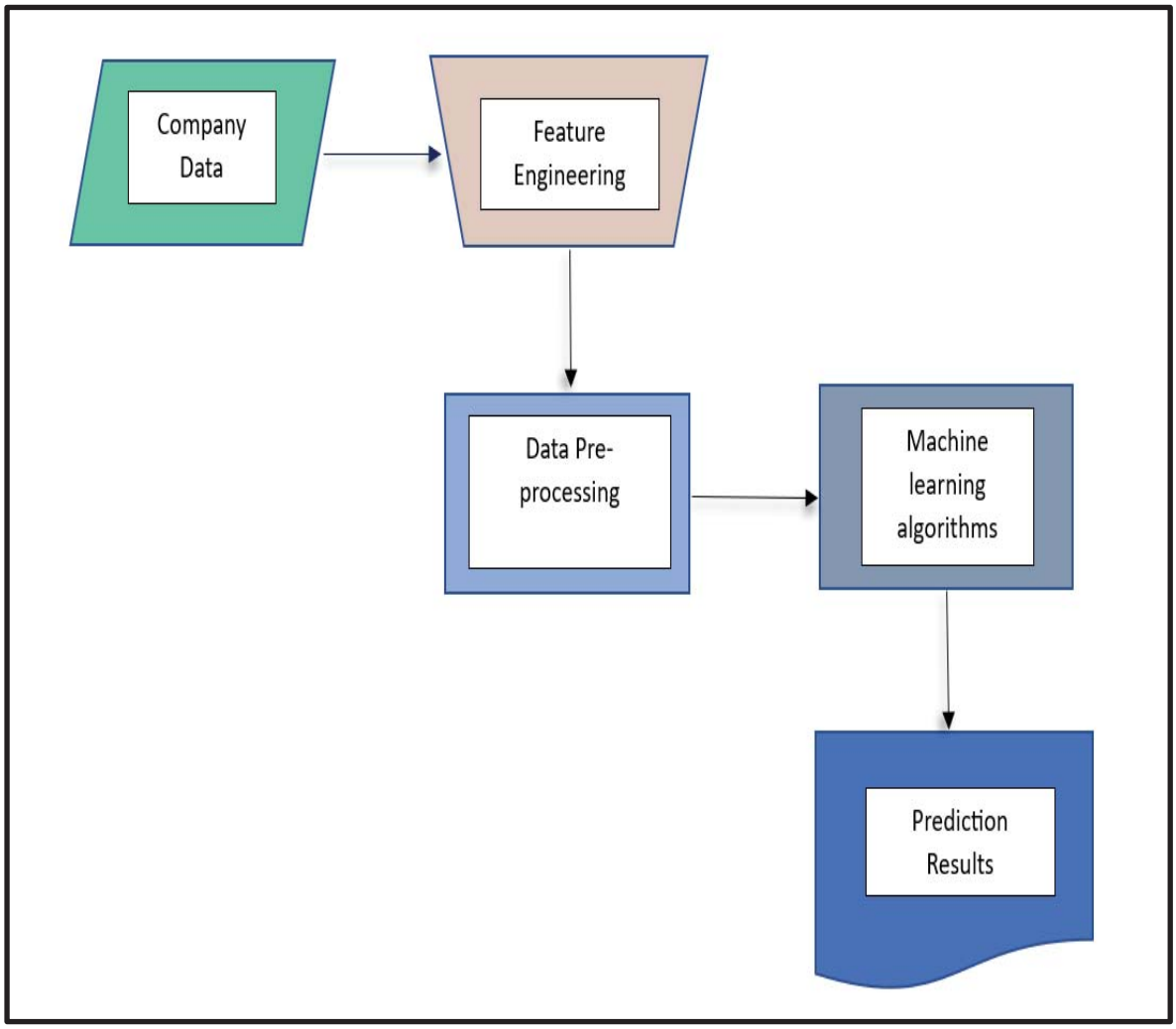

Figure 1: Basic ML Pipeline

Table 1: Churn- Prediction Methods: Comparison WRF: Weighted Random Forest, SVM: Support Vector Machine, FE_RF&T: based on Transduction and Random Forest, a feature extraction technique, OOPM: Orientation Ordering Pruning Methodology, AUC: Area Under the Curve, CFS: correlation-based feature selection

AUTHOR	DATASET	METHOD	ADVANTAG	DISADVANTAG	OUTCOME/

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

			ES	ES	FINDINGS
J. Burez et al, [1]	Six real-world, exclusive European churn modelling datasets	Gradient boosting technology, random and sophisticated undersampling, and WRF	Improves churn prediction accuracy.	The overall performance is reduced by the individual issues.	The findings demonstrate that under-sampling can increase prediction accuracy, particularly when AUC is used as a measurement. Weighted random forests outperform CUBE for cost-sensitive learner performance
Veronikha Effendy et al, [2]	Categorical type churn data	Combined sampling with WRF	High F-measure and prediction accuracy levels. Resolves imbalanced data issues.	The most common under-sampling technique is used.	In addition to selecting the correct target (by raising the value of the top decile), more research is required to

					establish which characteristics are most likely to cause customers to churn.
G. Ganesh-Sundarkumar- et al, [3]	Insurance dataset	Decision tree and one-class SVM undersampling	High precision and a simplified system	Detecting fraud is a more practical application than predicting turnover	AUC was significantly improved when undersampling using the sigmoid kernel compared to other methods, whereas undersampling using the radial basis kernel produced high AUC.
Qiu Yihui et al, [4]	Chie Mobile Communication's system for conducting business	FE RF&T feature extraction and OOPM feature	Incredibly precise churn forecast eliminates unneeded data.	Application requirements are only met based on test sample distribution data.	OOPM and FE_RF&T methods improve learning

	analyses.	selection			machine performance .
Qiuhua Shen et al, [5]	European telecommunications company data	Based on the complementary fusion of multiple-layer information, churn prediction	Improved dimensionality reduction and high prediction	The problem of unbalanced data recurs when feature selection is insufficient.	The "churn" prediction aim can be effectively reached with INTER, CMP, and CUSP. The experimental results showed that this method is better than CFS
Sebastián Maldonado et al, [6]	UCI-Telecom, Operator 1, Cell2Cell	Profit based SVM	Improved accuracy with a business focus.	Regulatory requirements are not met in SVM	According to experimental findings, our models perform better than traditional feature selection strategies in terms of business-related

					goals.
Xia Guo-en et al, [7]	UCI data & Home telecommunication carry dataset	SVM	Better accuracy even when there are many attributes present, a high churn rate, etc.	Inappropriate choice of kernel function and weights. There is a high dimensionality issue.	Customer churn was predicted using SVM on structural risk reduction, which had the highest accuracy, hit rate, covering rate, and lift coefficient.
Hsiu-Yu Liao et al, [13]	Roomi dataset	Hybrid classification with combined features	High forecast accuracy in a short amount of time.	There is a multi-objective issue.	Under various classification techniques, hybrid customer churn prediction performs well. It shows how virtual world platform providers

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

					can evaluate useractivity, neighboring activity, and neighborhood energy.

III. CHURN-PREDICTION METHODOLOGIES: COMPARISON

The methods examined are enumerated and contrasted in this part based on their benefits and drawbacks. Table 1 presents the comparisons. The numerous methodologies can be understood much more easily thanks to this table, which also helps the readers grasp the goal of the study. The table shows that hybrid approaches, such as hybrid machine learning offer high accuracy in churn prediction. SVM, ANN, SOM, and other hybrid models offer great accuracy with less complicated computations.

IV. CONCLUSION

The issue of client churn and the advantages of foreseeing attrition are first discussed in the context of telecom enterprises. The information on the datasets being examined and a description of the most significant churn prediction techniques currently in use are provided in the table above. Our review's emphasis on accurately anticipating churn as well as on the causes of churn and the shortcomings of current approaches is its most crucial component. All these processes aim to predict client turnover, with some utilizing direct machine learning approaches and others utilizing indirect methods to improve data pre-processing and feature selection strategies. Based on these findings, it can be said that hybrid techniques, as opposed to only one algorithm, produce the most accurate churn predictions. We are motivated to create a hybrid churn- prediction model of our own in the future given the wider breadth of the churn prediction research.

REFERENCES

[1] J. Burez and Dirk Van den Poel. "Handling class imbalance in customer churn prediction." Expert Systems with Applications, vol. 36, no. 3, pp. 4626-4636, 2009.

[2] Veronikha Effendy and ZK AbdurahmanBaizal. "Handling imbalanced data in customer churn prediction using combined sampling and weighted random forest." In 2014 2nd International Conference on Information andCommunication Technology (ICoICT), pp. 325-330. IEEE, 2014.

[3] G. Ganesh Sundarkumar, Vadlamani Ravi, and V. Siddeshwar. "One-class support vector machine based undersampling: Application to churn prediction and insurance fraud detection." In 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-7. IEEE, 2015.

[4] Qiu Yihui, and Zhang Chiyu. "Research of indicator system in customer churn prediction for telecom industry." In 2016 11th International Conference on Computer Science & Education (ICCSE), pp. 123-130. IEEE, 2016.

[5] Qiuhua Shen, Hong Li, Qin Liao, Wei Zhang, and KoneKalilou. "Improving churn prediction in telecommunications using complementary fusion of multilayer features based on factorization and construction." The 26th Chinese Control and Decision Conference (2014 CCDC), pp. 2250-2255. IEEE, 2014.

[6] Sebastián Maldonado, Álvaro Flores, Thomas Verbraken, Bart Baesens, and Richard Weber. "Profitbased feature selection using support vector machines– General framework and an application for customer retention." Applied Soft Computing, vol. 35, pp. 740- 748, 2015.

[7] Xia Guo-en and Wei-dong Jin. "Model of customer churn prediction on support vector machine." Systems Engineering-Theory & Practice, vol. 28, no. 1, pp. 71-77, 2008.

[8] María Óskarsdóttir, Cristian Bravo, Wouter Verbeke, Carlos Sarraute, Bart Baesens, and Jan Vanathien. "A comparative study of social network classifiers for predicting churn in the telecommunication industry." 2016.

[9] Umayaparvathi, V., and K. Iyakutti. "A Survey on Customer Churn Prediction in Telecom Industry: Datasets, Methods and Metrics." International Research Journal of Engineering and Technology (IRJET), vol 04, no.4, pp. 1065-1070, 2016.

[10] Yaya Xie, Xiu Li, E. W. T. Ngai, and Weiyun Ying. "Customer churn prediction using improved balanced random forests." Expert Systems with Applications, vol. 36, no. 3, pp. 5445-5449, 2009.

[11] Ionuţ Brânduşoiu, Gavril Toderean, and HoriaBeleiu. "Methods for churn prediction in the pre-paid mobile telecommunications industry." In 2016 International Conference on Communications (COMM), pp. 97-100. IEEE, 2016.

[12] Preeti K. Dalvi, Siddhi K. Khandge, Ashish Deomore, Aditya Bankar, and V. A. Kanade. "Analysis of customer churn prediction in telecom industry using decision trees and logistic regression." In Symposium on Colossal Data Analysis and Networking (CDAN), pp. 1-4. IEEE, 2016.

[13] Hsiu-Yu Liao, Kuan-Yu Chen, Duen-Ren Liu, and YiLing Chiu. "Customer Churn Prediction in Virtual Worlds." In 2015 IIAI 4th International Congress on Advanced Applied Informatics (IIAI-AAI), pp. 115-120. IEEE, 2015