Customer Churn Prediction in Telecom Sector: A Survey and way a head

Ibrahim AlShourbaji, Na Helian, Yi Sun, Mohammed Alhameed

Abstract—The telecommunication (telecom)industry is a highly technological domain has rapidly developed over the previous decades as a result of the commercial success in mobile communication and the internet. Due to the strong competition in the telecom industry market, companies use a business strategy to better understand their customers' needs and measure their satisfaction. This helps telecom companies to improve their retention power and reduces the probability to churn. Knowing the reasons behind customer churn and the use of Machine Learning (ML) approaches for analyzing customers' information can be of great value for churn management. This paper aims to study the importance of Customer Churn Prediction (CCP) and recent research in the field of CCP. Challenges and open issues that need further research and development to CCP in the telecom sector are explored.

Index Terms— Customer churn, prediction, machine learning, churn management, telecom

1 INTRODUCTION

Customers seek good service quality and competitive pricing factors in telecom sector. However, when these factors are missing, then they can easily leave to another competitor in the market [1]. This has led telecom to offer some incentives to customers to encourage them to stay [2].

The movement of customers (i.e., subscribers) from one service provider or carrier to another is called customer churn, It has been recognized that long-standing consumers are more lucrative in the long term, as new clients are engrossed by persuasive offers and incline to switch to an alternative competitor in the market at the moment they obtain a better concession [3-7], and therefore it is vital for companies to consider churn management as a part of Companies use CRM as a strategy to modify their process management, to improve their revenues and to find new approaches by primarily focusing on customers' needs to avoid losing them rather than a product [13, 14]. These specifics have led competitive companies to capitalize on CRM to up-hold their customers, and thus helping to increase customer strength. Figure 1 shows the main sections of CRM [15].

- Collaborative CRM: It aims to establish customized relationships with customers using several ways such as emails, telephone, websites, call centers, face-to-face contact, etc.
- Operative CRM: This type offers services for the organizations to increase the efficiency of CRM processes
- Analytical CRM: focuses on data collection and analysis to help the management build strategic decisions and plan for the future.

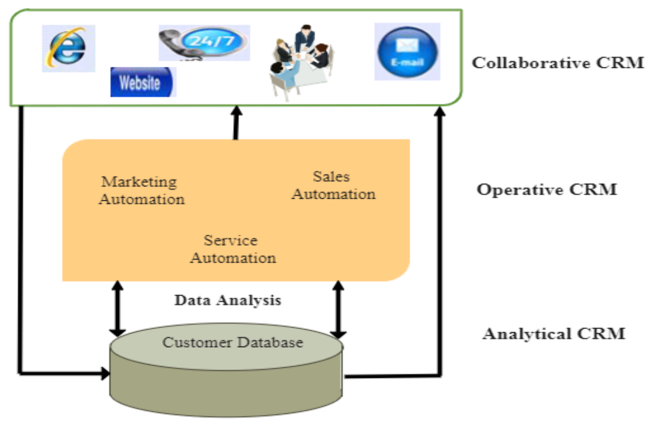

Fig 1.CRM areas

The data of customers are stored in such CRM systems which can then be transformed into valuable information with the help of ML techniques which aid telecom companies to formulate new polices, develop campaigns for existing clients and figure out the main reasons behind customer churn. In this way, companies can easily observe their customer's behavior from time to time and manage them effectively. Therefore, ML approaches are needed in telecom sectors which remain the corner-stone of customer churn control and can play a fundamental role in decreasing the probability of churners.
Due to the increased amount of data collection, organization and companies can store vast amount of data and information using several types of storage technologies at low cost. However, the challenge is to analyze, summarize and discover knowledge from these stored data. ML and statistics aiming at automatically discovering useful information and identifying hidden patterns in large data warehouses. ML involves few phases from raw data collection to some of the interesting patterns and this process includes data cleaning, transformation, selection and evaluation
It has been reported that attracting new customers' in the telecom sector, costs six times more than retaining existing

- *Ibrahim AlShourbaji, Na Helian and Yi Sun, School of Computer Science, University of Hertfordshire, Hatfield, U.K,*
- *Ibrahim AlShourbaji, Department of Computer and Network Engineering, Jazan University, 82822-6649 Jazan, Kingdom of Saudi*
- Mohammed Alhameed, *Department of Computer Science, Jazan University, 82822-6649 Jazan, Kingdom of Saudi Arabia*
- *Email: alshourbajiibrahim@gmail.com*

ones [7], and losing customers leads to a reduction in their sales [9]. Therefore, knowing the reasons for the customers to move to another service provider and the role of ML approaches can facilitate accurate segmentation of the customers and can yield significant returns to the telecom companies. This paper aims to provide the reader with the factors that can play a great effect on customer churn as well as the state-of- the- art of ML approaches applied in CCP in the telecom sector. Also the challenges and open areas that need further research in the application of CCP.

The rest of the paper is structured as follows: Section 2 provides the importance of CCP in the telecom sector. Section 3 presents the background literature from both the business domain knowledge and the computer science domain. Section 4 discusses the challenges and open issues; and finally, section 5 concludes this paper.

2 The Importance of CCP in Telecom

Information and Communication Technology (ICT) has grown and developed rapidly during the last decades, specifically in the mobile industry which represents the largest ICT market as a result of the appearance of the internet and the commercial success in the mobile communication market. Before 1999, the Internet was regarded as a fancy tool that only professionals, computer survey users and “nerds” could play with [16]. During that time, the number of Internet users was less than 5% of the population worldwide globally. Recently, ITU estimates the number of internet subscribers at the end of 2019 to reach 53.6 percent as shown in figure 2.

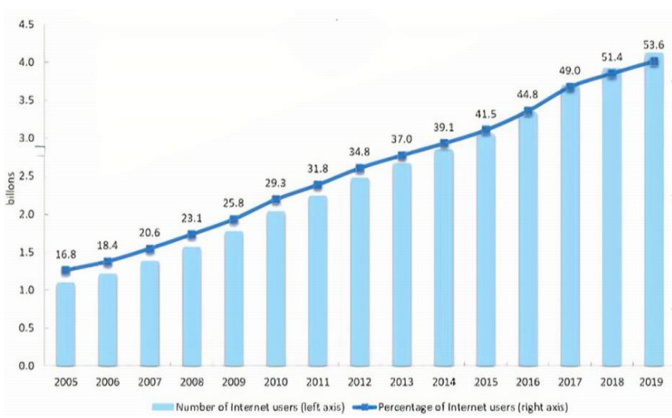

Fig 2. Subscribers using the internet 2005- 2019

By the end of 2020, the number of mobile phone users is expected to reach 6.918 billion, which is over 84% of the population globally [17]. However, the ICT market, particularly the telecom industry, has reached market saturation and the average annual churn rate reaches between 10 - 67% monthly due to the strong competition between service providers to attract new customers [18, 19].

Commercial companies in general and telecom companies in particular are considered as one of the top sectors that suffer from customer churning [20]. This means a company could lose approximately half of its customers and could result in a drop in its profits. Furthermore, research works from different countries such as Nigeria [21], India [22], Kenya [23], Indonesia [24] and Ghana [25] acknowledged the existence of the problem of churn in telecom companies.

Churn management is an essential concept in CRM; it manages the most fundamental aspects that may change the customers' behavior such as price, service quality, company’s reputation and effective advertising competition. Offering retention incentives is the primary way to reduce customer churn [26]. Telecom companies use different strategies for churn management and retention: offering incentives to all the clients without determining which customers to target [27], or utilizing customers’ transactional data to develop predictive models to specify the customers who are likely to defect in advance [28]. Once specified, the telecom company could target these customers with incentives to encourage them to stay [29].These incentives can take several forms such as promotions, discounts and free calls .etc.

Price reduction can be considered as the main market method in the telecom sector to retain and attract customers who are willing to churn [30, 31]. Free calls promotion in which the customers pay an equivalent amount of 12 or 15$ monthly and get free calls on the same network have become a common and result in affordable telecom service. This encouraged the existing subscribers to move to their calling plans and join their promotion because they have the choice to easily move and join another carrier at any time by purchasing a new SIM card which may leave the customers to have 3 or 4 different phone lines for him to use the one with the best promotion.

An effective deployment of quality systems in the telecom sector is vitally important to enable their companies attract, obtain and retain customers. It is certainly cost-effective to maintain existing customers than obtaining new ones [32].Therefore; the companies save millions of dollars by investing in customer churn analysis and prediction [33]. The companies through CRM work by ranking the margin of the propensity of customers to defect and facilitate their marketing team to provide incentives to those highly ranked. This saves the companies from losing the reputation and image dents of their brands and services. The telecom services are now customer-oriented or customer-centered [33].

The voice of the customer is incorporated in even a bit of production and service provisions to enhance satisfaction. Therefore, telecom firms should invest in churn prediction to maximize their profits, to reduce customer churn, and ensure that the needs as well as the behavior of their customers are being satisfied. Traditional data analysis techniques that focus on mining quantitative and statistical techniques are used for churn detection in the telecom industry [34]. The success of this model depends on observing the customers’ behavior with the help of experience and creating some rules to categorize a client as a churner or not. For instance, a telecom company could label the client as a churner when several calls are made by the client to the customer service. However, all of these rules are created using only intuition, experience and without the use of a scientific method, so the results may be below the expectations. In light of this, a powerful method is needed to effectively make reliable predictions and decisions than those who are just based on experience.

ML techniques have been broadly employed to model customer churn [35, 36]. This is because a churn is a rare event in a dataset and making accurate decisions requires creating models with high predictive performance [37]. Therefore, to reduce customer churn probability and developing an effective customer retention-program, the utilized predictive model should be accurate enough [38], or else, these systems will be useless when spending incentives on customers who will not churn. The true classification of churners and non-churners provides telecom companies enough time to build a specific campaign to decrease customer churn possibility and maximize their profitability from the retention campaign [39]. Figure 3 shows the main learning model phases:

i. Preparation phase
ii. Learning phase
iii. Performance and evaluation phase
iv. Decision phase

The customer churn data is prepared and transformed into an understandable format to make conceptual labeling in the first phase. Therefore, the labeled data contains training and testing in this phase. The main rule of the second phase is to prepare the input training data for building base classifiers with a learning algorithm as a base learner. The third phase chooses the most appreciate classifier for customer churn data. Finally, the decision-making phase tunes the learning rules to improve the accuracy of the prediction or detection of churn customers.

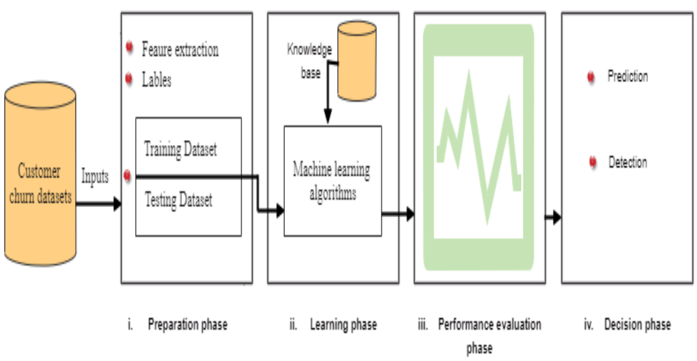

Fig 3. Learning system model

3 Literature Review

3.1 Business Domain knowledge

By looking at the market size in the telecom industry, with the intent of finding the necessary information to understand how large the churn problem is. Key findings include:

- The number of subscriber milestone in the telecom industry is expected to increase and reach more than 9 billion by the end of 2022 globally [40]
- 62.9% worldwide already owned a mobile phone in 2016 and it is expected to reach 67.1% in 2019 [41].
- 35% of the individuals using the Internet are young people aged 15-24, Least developed countries compared with 13% in developed countries and 23% worldwide [42].
- The following information is the average across the telecom industry:
- Monthly churn rate reaches between 10 - 67% [20]
- The average monthly revenue per business customer of Frontier communications reached 673.72 U.S. dollars in 2016 [43].
- The gross margin in the second quarter of 2017 is 80.38% [44].
- The customer lifetime is 52 months [45].
- A customer lifetime value is1782 U.S. dollars [45].
- The acquisition cost of a new customer in the Telecom sector is 315 U.S. dollars [46].

Customer lifetime value can be defined as a measure of how much profit can be generated over the lifetime of the customer.

Gross margin is a company's residual profit after selling a service or product and deduction the cost allied with its production and sale.

Acquisition cost refers to the cost of gaining new customers which involve persuading customers to purchase a company's services or products. It measures how much value customers bring to telecom companies and their businesses

Several reasons are presented in the work of [47] for when customers decide to stop using the service and why. The authors classified the causes of churn into three groups: controllable churn, uncontrollable churn and non pay/abuse. The controllable includes anything that is under the control of the company: Defecting to a competitor, response to poor service and the service price. Uncontrollable includes all the reasons that are outside the control of the company hands such as, death, illness and moving to a different country. The last group includes the causes that are related to nonpayment, abuse, theft of service or other causes in which the company made the churn decision for the customer and it is unclear whether any of the non-pay/abuse is controllable or not. A wide variety of factors play a great effect on churn in the telecom industry such as; income level, educational background, marital status, age, gender, geographical location, the effect of family and friends, cultural habits, service quality and price. In [48], the authors decided to investigate account length, international plan, voice mail plan, number of voice mail messages, total day minutes, total evening minutes, total night minutes, total international minutes, and number of calls to customer service factors. Some other works use income level, educational background, marital status and friends factors [49] and economic patterns: rate plans, tariffs and the promotion available from different service providers [50]. The author in [51] conducted research to analyze customer behavioral and demographic characteristics. The behavioral factors include rate plan (i.e., number of rate plan changes made with the carrier), handset changing frequency (i.e., number of handset changes made with the carrier), contract (Customer's service contract), rate plan suitability, customer tenure (i.e., number of months the customer stays with the service provider since service activation) and account status (still active or already churned at the end of the study period). The demographic factors include age, location (Western Canada or Eastern Canada) and language (English or French). The data were extracted from a

Canadian wireless carrier and 4896 residential customers were selected. The final results indicated that rate-plan suitability has an important role in customer churn among all the factors used and the customers who changed their call plans more than once, churned less. In another recent work, the authors in [52] demonstrated the lost customer-first behaviors and lifetime experience, the reasons behind defection and the nature of the win-back offer made to lost customers are all related to the likelihood of their reacquisition, lifetime duration, and the lifetime profitability per month in a US telecom products and services company. They include six sets of variables:

- First behaviors and lifetime experience (a member of referrals, number of complains and service recovery).
- Defection behavior (price-related reasons, service-related reasons, or price and service=related dummy and time of defection).
- Win back offer nature (price discount, service upgrade and price discount and service upgrade),
- Interaction between reasons for defection and win back offers (price-related defection X price discount offer, service-related defection X service upgrade offer).
- First-lifetime marketing contacts (frequency of phone calls, e-mails sent and direct mails).
- Demographics and first life control variables (age, gender, income, household size, education level, tenure, revenue, level of service plan and cross-buy).

Their empirical results indicated that referral and complain behaviors are important pointers for the quality of the first-lifetime experience and how the customers who have positive first-lifetime experiences were more likely to accept a win-back offer. There are other factors and their complex relationships affect customer churn such as service quality which is a combination of features such as network coverage, signal strength, voice quality, customer service is that provided by the service provider to the customer. This factor has a direct influence on encouraging customers to switch to another service provider as confirmed in the work of [53, 54]. The authors in [55] surveyed 196 respondents in the java west area to examine customer value and service quality factors to find their impact on controlling customer churn. They concluded that service quality has a positive relationship in controlling customer churn and also affect customer value positively. Service usage, switching cost, customer dissatisfaction and demographic factors play a very important role for a customer to switch to another service provider [56], when the authors used a dataset contains 1,000,000 records and 42 factors collected from a telecom company in the US.

***Table 1** summarizes the list of literature works and various factors used by researchers in their works.*

Works	Customer churn factors
(Braun & Schweidel, 2011, [47])	Controllable churn, uncontrollable churn and non pay/abuse
(Antipov & Pokryshevskaya, 2010, [48])	Account length, international plan, voice mail plan, number of voice mail messages, total day minutes, total evening minutes, total night minutes, total international minutes, and number of calls to customer service
(Wong, 2011, [49])	Income level, educational background, marital status and friends
(Ranaweera, 2007, [50])	Economic patterns: rate plans, tariffs and the promotion available from different service providers
(Wong, 2011, [51])	Behavioral: Rate plan, handset changing frequency, contract, rate plan suitability, customer tenure and account status, Customer demographic information: Age, location and language
(Kumar et al., 2015, [52])	First behaviors and lifetime experience, defection behavior, win back offer nature, the interaction between reasons for defection and win-back offers, First-lifetime marketing contacts, and customer demographics
(Cronin et al., 2000, [53]) (Al-Rousan et al., 2010, [54])	Service quality
(Marwanto & Komaladewi, 2017, [55])	Customer value and service quality
(Al-Mashraie et al., 2020, [56])	Customer demographic information: age and gender, number of dripped calls, number of service calls and geographical area

.

3.2 Computer Science Domain

Developing a model that accurately predicts customer churn could have several managerial and financial implications for telecom companies. The correct classification of a customer as churner and non-churner can reduce misclassification costs such as cost of incentives and retention rate in real-world decision making, the assumption of equal miss classification, the default operating mode for many classifiers, is most likely violated. Customers' churn has significant negative managerial and financial results on the company retention strategies. For example, a company could lose direct contact with the client, and therefore it cannot sell its additional products to that customer. Also, when an incentive sent to a non-churner client rather than a real one, it means that the real churner has not received the incentive to encourage him/her to stay or the incentive has received by a client who was not supposed to receive it and thus the marketing budget for the company could be negatively affected [57, 58]. A worldwide range of studies has applied ML techniques for CCP modeling during the last decade. Table 2 provides an overview of previous works on the use of ML techniques for modeling CCP in the telecom industry. Some main reference papers in recent literature, along with their titles, the used modeling techniques, the dataset names and their characteristics, the number of records and variables and whether the datasets are Private (*) or public (#), the applied evaluation measures, the validation method and research outcomes are summarized in the table.

Table 2. *Overview of previous works on churn prediction using ML approaches in the telecom industry*

works	Title of paper	Used Techniques	Dataset-#rec. - #var. - Private (*) or public (#)	Measures- var. selection - sampling - validation	Outcomes
(Idris et al., 2012, [59])	Genetic programming and adaboosting based churn prediction for telecom	Adaboost style boosting – Artificial Neural Network (ANN) and Random Forest (RF)	Orange telecom (KDD Cup 2009 small) –50,000 rec. – 230 var. - (#) and Cell2Cell - 70831sub. - 75 var. - (#)	AUC, sensitivity, specificity - Genetic Programming (GP) - uniform numerical format -no validation	GP- AdaBoost based model offers higher accuracy than ANN and RF. The GP-AdaBoost achieved a prediction accuracy of 89% for Cell2Cell and 63% for the other dataset.
(Miguéis et al., 2013, [60])	Customer attrition in retailing: an application of multivariate adaptive regression splines	Logistic Regression (LR)) and Multivariate Adaptive Regression Splines (MARS)	European retail company – 130284 #rec. – 7 var. - (*)	AUC, top-percentile lift - stepwise forward and stepwise backward – no sampling – 10- fold cross validation	MARS achieved better results when the whole set of variables are used. However, the LR outperforms MARS when variable selection procedures are applied.
(Brandusoiu &Toderean, 2013, [61])	Churn prediction in the telecommunications sector using support vector machines (SVM)	Support Vector Machine (SVM) with four kernel functions: Radial Basis Function kernel (RBF), linear kernel (LIN), Polynomial kernel (POLY) and sigmoid kernel (SIG)	UCI Repository, University of California - 3333 rec. – 21 var. -(#)	Recall, Specificity, Precision, Accuracy, Misclassification and F-measure – no var. selection – no sampling – no validation	SVM model using polynomial kernel (SVM-POLY) showed superiority over other SVM models.
(Lemmens & Gupta, 2013, [32])	Managing churn to maximize profits	gain/loss matrix and gradient boosting	Teradata Center at Duke University - -10,000 rec. – 171 var. – (*)	Misclassification - PCA –oversampling – F scores	The results indicated that improvements are achieved by using gain/loss matrix and gradient boosting approach for companies with no additional implementation cost.
(Keramati et al., 2014, [62])	Improved churn prediction in telecommunication industry using data mining techniques	Decision trees, ANN, K-Nearest Neighbors and SVM	Iranian mobile company – 3150 rec. – 11 var. - (*)	Accuracy, precision, recall, F –score – exhaustive recombination of variables - training-	ANN shown higher accuracy than Decision trees, K-Nearest Neighbors and SVM.

Table 2. *Overview of previous works on churn prediction using ML approaches in the telecom industry*

(Chen et al., 2015, [63])	Predicting customer churn from valuable B2B customers in the logistics industry: a case study	LR, Decision tree (C4.5), ANN (multilayer perceptron, (MLP)) and SVM	Logistics company in Taiwan - 69,170rec. - 18 var. - (*)	Accuracy, precision, recall, F1) - no var. selection - random re-sampling – 10-fold cross validation	C 4.5 outperformed the other models and. the most significant variables for customer churn: Length, recency and monetary.
(Vafeiadis et al., 2015, [64])	A comparison of ML techniques for CCP	SVM-poly, Decision tree, ANN, Naïve Bayes, Regression Analysis,	UCI Repository, University of California - 3333 rec. – 21 var. -(#)	Precision, Accuracy, Recall and F-measure - no var. selection – no sampling – monte carlo based cross validation	SVM-poly using AdaBoost obtained 97% accuracy and F-measure over 84%.
(Zhang et al., 2015 [65])	Profit Maximization Analysis Based on Data Mining and the Exponential Retention Model Assumption with Respect to Customer Churn Problems	DT and Regression	Guangxi Mobile Communication Company in China - 40,000 rec. – 127 var. – (*)	ROC , normalized profit - no var. selection – no sampling– no validation	The relationship between profit and retention is good when the prediction algorithm sufficiently good, when the capability of retention is good enough, the relationship of profit and retention is convex and when both prediction algorithm and retention capability are not effective enough, the operators should not take any actions.
(Hassouna et al, 2016, [66])	Customer Churn in Mobile Markets A Comparison of Techniques	Decision tree (CART, C 5.0, CHAID) and LR	Two UK mobile telecom operator data warehouse - 15,519 and 19, 919 rec. – 17 var.- (*)	AUC, Receiver Operating Characteristic (ROC), top decile, accuracy - choosing the most importantvar. - no validation	C 5.0model is better than the LR model for CCP.
(Umayaparvathi&Iyakutti, 2016, [67])	Attribute selection and CCP in telecom industry	Gradient Boosting (GB), DT, SVM, RF, K-Nearest Neighbour, Ridge Regression and LR	Cell2Cell - 70831sub. - 75 var. - (#) and CrowdAnalytix – 3333 rec. – 20 var. - (#)	Accuracy, Precession and Recall, F1-score - Brute force - no sampling - 10-fold cross validation	GB model has a higher performance than other techniques and six attributes: day minutes, voice, mail plan, night charge, international calls, evening calls, and day calls minutes have upmost importance towards churn prediction in the Cell2Cell dataset.
(Brânduşoiu et al., 2016, [68])	Methods for churn prediction in the pre-paid mobile telecommunications industry	Neural networks, SVM and Bayesian networks.	UCI Repository, University of California - 3333 rec. – 21 var. - (#)	Confusion matrix, gain measure, ROC - Principal Component Analysis (PCA) –oversampling – no validation	SVM achieved better performance compared to other techniques
(Coussement et al, 2017, [69])	A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry	CART, Bayesian network, J4.8 decision tree, MLP, Naive Bayes, RF, SVM with RBF kernel function and Stochastic gradient boosting	Large European mobile telecommunication provider - 30, 104 rec. – 956 var. - (*)	AUC, top decile lift (TDL) - correlation-based var.– no sampling – no validation	Data preparation treatment (DPT) improves prediction performance. Logistic regression- DPT approach outperformed the empirical methods remarkably.

Table 2. *Overview of previous works on churn prediction using ML approaches in the telecom industry*

(Prashanth et al, 2017, [70])	High Accuracy Predictive Modeling for CCP in Telecom Industry	Linear: LR, non-linear: RF, Deep Learning: Deep Neural Network, Deep Belief Networks and Recurrent Neural Networks	Asian telecom service provider – 337817 rec. – 36 var. - (*)	Accuracy, Sensitivity, AUC, Specificity- no var. selection - no sampling - 5-fold cross validation	Non-linear techniques performed better than the linear and both RF and deep learning gave comparable performance,
(Amin et al, 2017, [71])	CCP in the telecommunication sector using a rough set approach	Rough Set Theory (RST), Exhaustive Algorithm (EA), Genetic Algorithm (GA), Covering Algorithm (CA) and LEM2 algorithm (LA)	UCI Repository, University of California - 3333 rec. – 21 var. -(#)	Recall, Specificity, Precision, Accuracy, Misclassification, F-measure - Information Gain Attribute Evaluation - random re-sampling – K-fold cross validation	RST-GA based model showed satisfactory results for extracting implicit knowledge.
(Azeem et al., 2017, [72])	A churn prediction model for prepaid customers in telecom using fuzzy classifiers	Neural Network, Linear regression, C4.5, SVM, AdaBoost, Gradient Boosting, RF and Fuzzy classifiers: Fuzzy NN, VQNN, OWANN and Fuzzy Rough NN	telecom company in South Asia – 600,000 rec. – 722 var. – (*)	Recall, Precision, ROC, AUC - domain knowledge – oversampling – training-test split	Fuzzy classifiers shown superior performance compared to other used models
(Zhu et al., 2017, [73])	Benchmarking sampling techniques for imbalance learning in churn prediction	LR, Decision Tree (C4.5), SVM and RF	Chile- 5300 rec. – 41 var. – (*), Duke current– 51, 306 rec. – 41 var. – (*), Dukefuture – 100, 462 rec. – 173 var. – (*), Korean1 K1 Operator East Asia - 2019 rec. – 10 var. – (*), Korean2 K2 Operator East Asia -2941 rec. – 14 var. – (*), Korean3 K3 Operator East Asia -5990 rec. – 36 var. – (*), Korean4 K4 Operator East Asia -2183 rec. – 9 var. – (*), Korean5 K5 Operator East Asia - 26224 rec. – 11 var. – (*), Tele1 Operator Europe - 4350 rec. – 87 var. – (*), UCI Repository, - 3333 rec. – 19 var. -(#), KDD cup KDDcup - 50000 rec. – 231 var. -(#)	AUC, Maximum Profit (MP), top-decile lift - Fisher score - ADASYN, Borderline-SMOTE, CLUS, MWMOTE, SMOTE, SMOTE–ENN, SMOTE–Tomek – 5 × 2 cross-validation	Sampling approaches power lies in the used evaluation metrics as well as the classifiers.

Table 2. *Overview of previous works on churn prediction using ML approaches in the telecom industry*

(Effendy et al., 2014, [74])	Handling Imbalanced Data in CCP Using Combined Sampling and Weighted RF	Weighted RF(WRF)	telecom company in Indonesia – 48,384 rec. – 24 var. –	Recall, Precision and F-measure – no var. selection - combination of undersampling and SMOTE - 10-fold cross validation	The combined sampling techniques able to help WRF algorithm to achieve better performance and predict the churn effectively
Gui, 2017, [75]	Analysis of imbalanced data set problem: The case of churn prediction for telecommunication	RF	telecom company, 450.000 rec. - 22var. (*)	Recall, Precision F-measure, Cost, and Accuracy – RF, Relative Weight (RW), Scandalized Regression Coefficient (SRC) - oversampling, under sampling, SMOTE - no validation	SRC combined with SMOTE method achieved the best results
(De Caigny et al., 2018, [76])	A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees	Decision trees and LR	Financial services, 117, 808 rec - 237 var- (*), Retail, 32, 371 rec- 47 var- (*), DIY 3, 827 rec - 16 var- - (*), Newspaper, 427, 833 rec- 165 var- (*), Telecom, 71, 074 rec- 87 var - (*), Financial services, 102, 279 rec- 138 var - (*), Telecom, 47, 761 rec- 43 var- (*),,Telecom, 50,000 rec – 303 var (*), Financial services, 631, 627 rec – 232 var - (*), Financial services, 573, 895 rec – 232 var - (*), Financial services, 398, 087rec – 232 var- (*), Financial services, 316, 578 rec – 232 var - (*), Financial services, 602, 575 rec – 232 var - (*), Energy, 20, 000 rec – 33 var -(*)	AUC, top decile lift	The hybrid model provided more accurate model than using its building blocks; decision trees and LR; as a standalone classification models
(Ullah et al., 2019, [77])	A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector	JPK, LR, MLP, Naïve Bayes, AdaBoostM1, attribute selected classifier, decision stump, RF, J48, random tree and LWL	Call Detail Records company in South Asian– 64,107 rec – 29 var.- (*), UCI Repository, - 3333 rec. – 19 var. -(#)	Accuracy, Recall, Precision and F-measure	RF performed better in terms of prediction of churners
(Jafari-Marandi et al., 2020, [78])	Inferring Machine Learning-Based Parameter Estimation for Telecom Churn Prediction	ANN, Self-organizing map	Telecom company in Iran- –3150 – 12 var.- (*),	Accuracy, Recall, Precision, F-measure, and misclassification error	The proposed profit-driven models proved to be effective for telecom churn prediction

4 CHALLENGES AND OPEN ISSUES

Different researchers have tested different factors and their significant effect on churn such examples of these factors include age, gender, rate plans, and minutes of use.etc. However, Customer often compares their service provider with other competitors in the market and churn to whoever they feel provides better service Researchers have found service quality [53, 54, 55] as the top factor for customer churn. Poor of customers care and slow response to their needs and complaints are another factor which can also play a vital role in increasing the probability of customers to change to another service provider [79]. Another factor that needs further attention is the customer's tenure (i.e., churn time) with the service provider and its relationship with the prediction of customer churn in the telecom sector as confirmed in the work of [80, 81].

It has been confirmed that a high correlation exists between churn risk and customer dissatisfaction and the role in the turn of the telecom companies becomes that of preventing those dissatisfied customers from chinning [82]. To tackle the problem of customer churn, telecom companies use different strategies for churn management. These tactics include: identifying churners first by using a predictive model, followed by targeting those customers with retention incentives to encourage them to stay [11]. Another approach focused on customers who have the highest probability of defection (e.g. top 10%) [34]. A recent approach focused on choosing the target customers based on the profit potential of each, the likelihood of churning, the number of customers the company decides to target and the incentive cost to maximize the overall return from the retention campaign [54, 55].

Despite intensive efforts, there is no general judgment exists on the performance of the predictive techniques for customers churn prediction in the telecom sector. Therefore, a review and benchmarking experiment need to be done to allow comparing the performance of a variety of classification techniques. All kinds of data have different attributes that might pose problems for ML techniques to extract the most crucial patterns in datasets due to class imbalanced in datasets. The classes whose number of instances are below the average number of instances per calls are termed as minority classes, while the instances that are above the average number of instances per class are termed as majority classes. Many studies are carried out comparisons on sampling techniques for handling class imbalance problem in the preprocessing phase. However, none of the literature reached to a final conclusion about which sampling technique is suitable for customers churn problem in telecom sector. .

Feature selection main goal is to select the most important features without changing the original data representation, and therefore, selecting a subset of the features relevant for the task to achieve the highest classification accuracy. Searching for the optimal subsets of features is a necessary preprocessing step in ML techniques. Variable or feature selection eliminates irrelevant features and obtains the best feature subsets that play a vital role in the models' accuracy and their final results. For instance, Migueis et al., 2013, [60], evaluated the performance of two techniques: MARS and logistic regression to model customer attrition. The results showed that MARS performed better than Logistic regression when the whole variables in the dataset were used, but when the authors applied variable selection, the logistic regression, achieved higher accuracy. However, none of the previous works confirmed which ML techniques can perform the best for feature selection.

To evaluate the performance of a classifier, it is essential to specify at the beginning which evaluation measures should be used to fix the optimization objective for the entire analysis. Churn prediction problem is considered as a binary classification task because the outcome has only two possible values (churner or nonchurner), accuracy, recall, and precision measures are often used to evaluate the classification quality of binary classifiers [83]. Recall measures the proportion of positive cases (churners) that are correctly classified. Precision measures the proportion of cases classified as churners that are correctly churners. In other works, the authors confirmed that the use of the confusion matrix represents the primary source for accurate estimation of churn prediction in the telecom sector [84].

Despite the intensive research efforts, there is still no consensus exists on the validation of results and far from agreeing on any standards. Therefore, it is important to verify how the model performs. In ML, validation strategies are used to validate a model's performance, and generalized their outcomes. One of the useful and popular mechanisms for validating a customer churn model performance and its accuracy is the cross-validation technique [29].

5 CONCLUSIONS

Creating and delivering superior customer value is considered one of the corner-stones in the telecom market. Customer churn is a term that is developed in the telecom sector which indicates the movement of customers from their service provider to another competitor in the market. Knowing the reasons for churning and predicting which customers are about to churn can yield a significant return in the profitability of the telecom companies. This paper gives a complete survey on the importance of customer churn, reasons behind customers churn and the state-of- the- art of ML approaches applied in CCP. This paper briefly describes the previous efforts in the field of CCP. This paper briefly discusses research challenges and future directions of the CCP in telecom sectors.

REFERENCES

[1] *Owczarczuk, M. (2010), Churn models for prepaid customers in the cellular telecommunication industry using large data marts, Expert Systems with Applications, 37(6) pp.4710-4712.*

[2] Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining *techniques for customer relationship management. Technology in society, 24(4), 483-502.*

[3] *Coussement, K., & De Bock, K. W. (2013). Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning. Journal of Business Research, 66(9), 1629-1636.*

[4] [4] Statista (2016), Average Monthly Churn Rate for Wireless Carriers in the United States from 1st Quarter 2013 to 1st Quarter 2016,available at: https://www.statista.com/statistics/283511/average-monthly-churn-rate-top-wireless-carriers-us/, (Accessed 19, September 2020)

[5] Yu-Teng, C. (2015) 'Measuring the impact of datamining on churn management', Internet Research: Electronic Networking Applications and Policy, Vol. 11, No. 5, pp. 375–387.

[6] Hashmi, N., Butt, N. A., & Iqbal, M. (2013). Customer Churn Prediction in Telecommunication A Decade Review and Classification. International Journal of Computer Science, 10(5), 271-282.

[7] Saradhi, V. V., &Palshikar, G. K. (2011). Employee churn prediction. Expert Systems with Applications, 38(3), 1999-2006.

[8] AlOmari, D., & Hassan, M. M. (2016, September). Predicting Telecommunication Customer Churn Using Data Mining Techniques.In International Conference on Internet and Distributed Computing Systems (pp. 167-178).Springer International Publishing.

[9] Coussement, K., & Van den Poel, D. (2013). Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers. Expert Systems with Applications, 36, 6127-6134.

[10] Risselada, H., Verhoef, P. C., &Bijmolt, T. H. (2010). Staying power of churn prediction models. Journal of Interactive Marketing, 24(3), 198-208.

[11] Coltman, T. (2007). Why build a customer relationship management capability?. The Journal of Strategic Information Systems, 16(3), 301-320.

[12] Ling, R., & Yen, D. C. (2001). Customer relationship management: An analysis framework and implementation strategies. Journal of computer information systems, 41(3), 82-97.

[13] Lin, C. S., Tzeng, G. H., & Chin, Y. C. (2011). Combined rough set theory and flow network graph to predict customer churn in credit card accounts. Expert Systems with Applications, 38(1), 8-15.

[14] Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., ...&Sriram, S. (2006). Modeling customer lifetime value. Journal of service research, 9(2), 139-155.

[15] Buttle, F. (2009). Customer relationship management: concepts and technologies. Routledge.

[16] ITU (2017), ICT Facts and Figures 2017, available at: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf (Accessed 19, September 2020).

[17] The Radicati Group, INC, a technology market research firm, Mobile Statistics Report 2014-2018, available at: https://www.radicati.com/wp/wp-content/uploads/2016/01/Mobile-Growth-Forecast-2016-2020-Executive-Summary.pdf, (Accessed 19, September 2020).

[18] Database Market institute (2017), Churn reduction in the telecom industry, available at: http://www.dbmarketing.com/telecom/churnreduction.html (Accessed Accessed 2, October 2020).

[19] Database Marketing Institute. Churn reduction in the telecom industry a valuable at:http://www.dbmarketing.com/2010/03/churn-reduction-in-the-telecom-industry/, (Accessed 2, October 2020).

[20] AT&T inc financial review 2019, Available at: https://investors.att.com/~/media/Files/A/ATT-IR/financial-reports/annual-reports/2019/complete-financial-review-2019.pdf , (Accessed 2, October 2020).

[21] Oghojafor, B., Mesike, G., Bakarea, R., Omoera, C., &Adeleke, I. (2012). Discriminant analysis of factors affecting telecoms customer churn. International Journal of Business Administration, 3(2), 59.

[22] Tenhunen, S. (2008). Mobile technology in the village: ICTs, culture, and social logistics in India. Journal of the Royal Anthropological Institute, 14(3), 515-534.

[23] Halim, J &Vucetic J (2015). Causes of Churn in the Wireless Telecommunication Industry in Kenya. DOI: 10.14355/ijmser.2016.0301.01

[24] Saefuddin, A., Setiabudi, N. A., &Achsani, N. A. (2011). The effect of overdispersion on regression based decision with application to churn analysis on Indonesian mobile phone industry.

[25] Sey, A. (2009). Exploring mobile phone-sharing practices in Ghana. info, 11(2), 66-78.

[26] Shaffer, Greg, and Z. John Zhang. "Competitive one-to-one promotions." Management Science 48, no. 9 (2002): 1143-1160.

[27] Hadden, J., Tiwari, A., Roy, R., &Ruta, D. (2007). Computer assisted customer churn management: State-of-the-art and future trends. Computers & Operations Research, 34(10), 2902-2917.

[28] Coltman, T. (2007). Why build a customer relationship management capability?. The Journal of Strategic Information Systems, 16(3), 301-320.

[29] Neslin, Scott A., Sunil Gupta, Wagner Kamakura, Junxiang Lu, and Charlotte Mason. "Defection detection: Measuring and understanding the predictive accuracy of customer churn models." Journal of marketing research 43, no. 2 (2006): 204-211.

[30] New vision daily, available at: https://www.newvision.co.ug/new_vision/news/1006093/telecom-price-war-sustainable-eur-mtn-boss, (Accessed 2, October 2020).

[31] Chen, Y., Zhang, G., Hu, D., & Fu, C. (2007). Customer segmentation based on survival character. Journal of intelligent manufacturing, 18(4), 513-517.

[32] Lemmens, A., & Gupta, S. (2013). Managing churn to maximize profits. Working Paper, http://www.hbs.edu/faculty/Publication%20Files/14-020_3553a2f4-8c7b-44e6-9711-f75dd56f624e.pdf

[33] Bit Refine group available at:https://bitrefine.group/aboutcompany/news/248-mobile-network-operators-combat-customer-churn-through-prediction-breakthrough, (Accessed 2, October 2020).

[34] Jiang, W., Au, T., &Tsui, K. L. (2007). A statistical process control approach to business activity monitoring. Iie Transactions, 39(3), 235-249.

[35] Wierenga, B. (2010). Marketing and artificial intelligence: Great opportunities, reluctant partners. In Marketing intelligent systems using soft computing (pp. 1-8). Springer Berlin Heidelberg.

[36] [36] Risselada, H., Verhoef, P. C., &Bijmolt, T. H. (2010). Staying power of churn prediction models. Journal of Interactive Marketing, 24(3), 198-208.

[37] [37] Kamakura, W., Mela, C. F., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., ...& Wedel, M. (2005). Choice models and customer relationship management. Marketing Letters, 16(3), 279-291.

[38] Coussement, K., & Van den Poel, D. (2008). Integrating the voice of customers through call center emails into a decision support system for churn prediction. Information & Management, 45(3), 164-174.

[39] Gordini, N., &Veglio, V. (2017). Customers churn prediction and marketing retention strategies. an application of support vector machines based on the auc parameter-

selection technique in b2b e-commerce industry. Industrial Marketing Management, 62, 100-107.

- [40] *Statista, a valuable at: https://www.statista.com/topics/1147/mobile-communications/, (Accessed 5, October 2020).*
- [41] *Statista, a valuable at:https://www.statista.com/statistics/470018/mobile-phone-user-penetration-worldwide/, (Accessed 5, October 2020)*
- [42] *ICT Facts and Figures, a valuable at: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf, (Accessed 5, October 2020)*
- [43] *Statista, a valuable at:https://www.statista.com/statistics/692056/frontier-communications-average-monthly-revenue-per-customer/, (Accessed 5, October 2020)*
- [44] *Seeking Alpha, Chart: Telecom Companies - Gross Profit Margins, a valuable at:https://seekingalpha.com/article/10168-chart-telecom-companies-gross-profit-margins, (Accessed 5, October 2020)*
- [45] *Churn in the telecom industry – identifying customers likely to churn and how to retain them.Aditya Kapoorm a valuable at:https://wp.nyu.edu/adityakapoor/2017/02/17/churn-in-the-telecom-industry-identifying-customers-likely-to-churn-and-how-to-retain-them/, (Accessed 5, October 2020).*
- [46] *Propeller a valuable at:https://www.propellercrm.com/blog/customer-acquisition-cost,, (Accessed 5, October 2020)*
- [47] *Braun, M., &Schweidel, D. A. (2011). Modeling customer lifetimes with multiple causes of churn. Marketing Science, 30(5), 881-902.*
- [48] *Antipov, E., & Pokryshevskaya, E. (2010). Applying CHAID for logistic regression diagnostics and classification accuracy improvement. Journal of Targeting, Measurement and Analysis for Marketing, 18(2), 109-117.*
- [49] *Wong, K. K. K. (2011). Getting what you paid for: Fighting wireless customer churn with rate plan optimization. Journal of Database Marketing & Customer Strategy Management, 18(2), 73-82.*
- [50] *Ranaweera, C. (2007). Are satisfied long-term customers more profitable? Evidence from the telecommunication sector. Journal of Targeting, Measurement and Analysis for Marketing, 15(2), 113-120.*
- [51] *Wong, K. K. K. (2011). Using Cox regression to model customer time to churn in the wireless telecommunications industry. Journal of Targeting, Measurement and Analysis for Marketing, 19(1), 37-43.*
- [52] *Kumar, V., Bhagwat, Y., & Zhang, X. A. (2015, May). Regaining "Lost" Customers: The Predictive Power of First-Lifetime Behavior, the Reason for Defection, and the Nature of the Win-Back Offer. American Marketing Association.*
- [53] *Lemmens, A., &Croux, C. (2006). Bagging and boosting classification trees to predict churn. Journal of Marketing Research, 43(2), 276-286.*
- [54] *Foster P, and Fawcett T. (2013). Data Science for Business: What You Need to Know About Data Mining and Data Analytic Thinking, O'Reilly Media, 2013.*
- [55] *Marwanto, S. T., & Komaladewi, R. (2017). How to restrain customer churn in telecommunication providers: study in west java Indonesia. Review of Integrative Business and Economics Research, 6, 51.*
- [56] *[56] Al-Mashraie, M., Chung, S. H., & Jeon, H. W. (2020). Customer switching behavior analysis in the telecommunication industry via push-pull-mooring framework: a machine learning approach. Computers & Industrial Engineering, 106476.*
- [57] *Coussement, K. (2014). Improving customer retention management through cost-sensitive learning. European Journal of Marketing, 48(3/4), 477-495.*
- [58] *] Gordini, N., &Veglio, V. (2017). Customers churn prediction and marketing retention strategies. an application of support vector machines based on the auc parameter-selection technique in b2b e-commerce industry. Industrial Marketing Management, 62, 100-107.*
- [59] *Idris, A., Khan, A., & Lee, Y. S. (2012, October). Genetic programming and adaboosting based churn prediction for telecom. In Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on (pp. 1328-1332).*
- [60] *Miguéis, V. L., Camanho, A., & e Cunha, J. F. (2013). Customer attrition in retailing: an application of multivariate adaptive regression splines. Expert Systems with Applications, 40(16), 6225-6232.*
- [61] *Brandusoiu, I., &Toderean, G. (2013). Churn prediction in the telecommunications sector using support vector machines. Margin, 1, x1.*
- [62] *Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., &Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. Applied Soft Computing, 24, 994-1012.*
- [63] *Chen, K., Hu, Y. H., & Hsieh, Y. C. (2015). Predicting customer churn from valuable B2B customers in the logistics industry: a case study. Information Systems and e-Business Management, 13(3), 475-494.*
- [64] *Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., &Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 55, 1-9.*
- [65] *Zhang, Z., Wang, R., Zheng, W., Lan, S., Liang, D., &Jin, H. (2015, November). Profit Maximization Analysis Based on Data Mining and the Exponential Retention Model Assumption with Respect to Customer Churn Problems. In Data Mining Workshop (ICDMW), 2015 IEEE International Conference on(pp. 1093-1097).*
- [66] *Hassouna, M., Tarhini, A., Elyas, T., &AbouTrab, M. S. (2016). Customer Churn in Mobile Markets A Comparison of Techniques. arXiv preprint arXiv:1607.07792.*
- [67] *Umayaparvathi, V., &Iyakutti, K. (2016, March). Attribute selection and Customer Churn Prediction in telecom industry. In Data Mining and Advanced Computing (SAPIENCE), International Conference on (pp. 84-90).IEEE.*
- [68] *Brânduşoiu, I., Toderean, G., &Beleiu, H. (2016, June). Methods for churn prediction in the pre-paid mobile telecommunications industry.In Communications (COMM), 2016 International Conference on (pp. 97-100).*
- [69] *Coussement, K., Lessmann, S., &Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. Decision Support Systems, 95, 27-36.*
- [70] *[70] Prashanth, R., Deepak, K., &Meher, A. K. (2017, July). High Accuracy Predictive Modelling for Customer Churn Prediction in Telecom Industry.In International Conference on Machine Learning and Data Mining in Pattern Recognition (pp. 391-402).Springer, Cham.*
- [71] *Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A., & Huang, K. (2017). Customer churn prediction in the telecommunication sector using a rough set approach. Neurocomputing, 237, 242-254.*
- [72] *[72] Azeem, M., Usman, M., & Fong, A. C. M. (2017). A churn prediction model for prepaid customers in telecom using fuzzy classifiers. Telecommunication Systems, 1-12.*

[73] Zhu, B., Baesens, B., &Backiel, A. (2017). Benchmarking sampling techniques for imbalance learning in churn prediction. Journal of the Operational Research Society.

[74] Effendy, V., & Baizal, Z. A. (2014, May). Handling imbalanced data in customer churn prediction using combined sampling and weighted random forest. In Information and Communication Technology (ICoICT), 2014 2nd International Conference on (pp. 325-330)..

[75] Gui, C. (2017). Analysis of imbalanced data set problem: The case of churn prediction for telecommunication. Artificial Intelligence Research, 6(2), 93.

[76] De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269(2), 760-772.

[77] Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., & Kim, S. W. (2019). A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector. IEEE Access, 7, 60134-60149.

[78] Jafari-Marandi, R., Denton, J., Idris, A., Smith, B. K., & Keramati, A. (2020). Optimum profit-driven churn decision making: innovative artificial neural networks in telecom industry. Neural Computing and Applications, 32(18), 14929-14962

[79] Sathish, M., Kumar, K. S., Naveen, K. J., & Jeevanantham, V. (2011). A study on consumer switching behaviour in cellular service provider: A study with reference to Chennai. Far East Journal of Psychology and Business, 2(2), 71-81.

[80] Mahajan, V., Misra, R., & Mahajan, R. (2017). Review on factors affecting customer churn in telecom sector. International Journal of Data Analysis Techniques and Strategies, 9(2), 122-144.

[81] Gopal, R. K., & Meher, S. K. (2008, May). Customer churn time prediction in mobile telecommunication industry using ordinal regression. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 884-889). Springer, Berlin, Heidelberg.

[82] Radcliffe, N. J., & Simpson, R. (2008). Identifying who can be saved and who will be driven away by retention activity. Journal of Telecommunications Management, 1(2).

[83] Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3), 4626-4636.

[84] Faris, H., Al-Shboul, B., & Ghatasheh, N. (2014, September). A genetic programming based framework for churn prediction in telecommunication industry. In International Conference on Computational Collective Intelligence (pp. 353-362). Springer, Cham.