Systematic Review of Customer Churn Prediction in the Telecom

Kamya Eria[1] and Booma Poolan Marikannan[2]
[1,2]Faculty of Computing, Engineering & Technology, Asia Pacific University of Technology & Innovation, 57000 Kuala Lumpur, Malaysia
[1]tp048982@mail.apu.edu.my, [2]dr.booma@apu.edu.my

Abstract- **The Telecommunications (telecom) Industry is saturated and marketing strategies are focusing on customer retention and churn prevention. Churning is when a customer stops using a company's service thereby opting for the next available service provider. This churn threat has led to various Customer Churn Prediction (CCP) studies and many models have been developed to predict possible churn cases for timely retention strategies. This review looks at the existing models in literature, using 30 selected CCP studies particularly from 2014 to 2017. Data preparation methods and churn prediction challenges have also been explored. This study reveals that Support Vector Machines, Naïve Bayes, Decision Trees and Neural Networks are the mostly used CCP techniques. Feature selection is the mostly used data preparation method followed by Normalization and Noise removal. Undersampling is the mostly preferred technique for handling telecom data class imbalances. Imbalanced, large and high dimensional datasets are the key challenges in telecom churn prediction. The issue of large telecom datasets is handled by use of new big data technologies such as Hadoop, Hbase and NoSQL (Not only Structured Query Language) though their adoption is still at a low rate.**

Index Terms- **Churn prediction, Customer churn, Data preparation, Telecommunication Introduction**

1. Introduction

Telecommunication (telecom) companies currently face a lot of competition due to the many industry players. This poses a serious threat among companies because customers have many options to switch to (Fei, Shuan, & Yan, 2017), which is termed as customer churn. Churning has attracted a lot of research from various scholars to aid timely detection of churners before they practically use their switching intentions (Umayaparvathi & Iyakutti, 2016). Timely detection of potential churners saves telecom companies from persuading costs that would be incurred to attract new customers or win back already churned customers. Such costs are 5 to 6 times higher than customer retention costs (Idris, Iftikhar, & Rehman, 2017; Zhu, Baesens, & vanden Broucke, 2017).

Many studies have revealed that high prices are among the top factors causing customer churn. Akmal (2017)'s qualitative study found out that high rates and bills are key factors causing churn. Mehwish, Zaffar and Sumaira (2017) also confirmed that network charges are among the top factors causing churn. Churning is also accelerated with good Mobile Number Portability (MNP) services (Adebiyi, Oyatoye and Amole 2016). Companies therefore have to try as much as possible to avoid customer churn cases because customers who leave a company have capabilities to influence members of their social groups to do the same (De Caigny, Coussement, & De Bock, 2018). Telecom companies deal with two kinds of customers namely Business to Consumer (B2C) and Business to Business (B2B) customers. B2C transactions provide telecom services to individual consumers while B2B transactions provide telecom services to other businesses.

In order to classify customers into potential churners and non-churners, Mitrović et al. (2017) confirmed that various Customer Churn Prediction (CCP) studies have been carried out and many models and techniques provided over time. However, many of these CCP studies initially did not consider the profit maximization objective which is the ultimate objective of any profit organization. Coussement, Lessmann and Verstraeten (2017) stated that companies should clearly distinguish non-profitable customers from those that are profitable. Profitable customers who are also potential churners are then given special attention with retention measures. As a result, retention campaigns-related costs are minimized because resources are utilized on the right and targeted customer group.

CCP relies on machine learning algorithms to develop models that classify telecom customers into churners and non-churners. Churn modelling includes three major steps after data collection and before model deployment; data preparation, model training and model evaluation (Umayaparvathi & Iyakutti, 2016). Data preparation is aimed at making data suitable for machine learning algorithms and model training. 50-80% of the data mining effort is put to data preparation because the quality of data affects the model performance results (Zhang, Zhang, & Yang, 2010). Data preparation is also responsible for removing any bias in the data through class balances and other randomization procedures. Missing value imputation, data cleaning, transformation and general exploration is done in this phase (Federico, 2014). Data transformations depend on the input data as well as the algorithms of choice (Olson & Delen, 2008). After the data has been prepared, it is spilt into a training and a validation set. The training set is used to train the models which are later tested with the validation set to evaluate their performance accuracy and reliability.

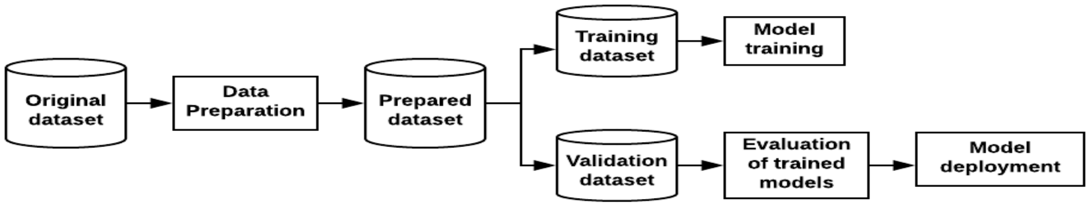

Figure 1: The churn prediction modelling process

This study aims at reviewing the available literature in the B2C context, observing how the various approaches to churn prediction have been applied shown in Figure 2. The study also looks at the data preparation methods employed to handle the huge telecom datasets. Data preparation aims at changing data into a form that makes it easy for churn prediction models to perform more accurately (Coussement, Lessmann and Verstraeten, 2017). The remainder of this study is organized as follows; the next section reports about the related works, followed by materials and methods used in the. Results are discussed and research findings are presented in the fourth and fifth sections respectively while conclusions are summarized in the sixth section.

2. Related Work

Hashmi, Butt and Iqbal (2013) carried out a systematic review on customer churn prediction in the telecom industry, considering studies carried out from 2002 to 2013. The author's selected 61 articles from the 834 initially availed in their primary search. In their study, it was found out that Decision Tree was the most used churn prediction technique. The study also revealed that the huge volumes, high dimensionality and imbalanced structure of telecom data is a big challenge for industry practitioners as they try to draw actionable insights from it. These issues give this study a foundation to consider literature from 2014 to 2017 and to further review literature as well as seeking solutions to the highlighted issues.

Several other studies have been carried out in form of surveys about CCP methods. Umayaparvathi and Iyakutti (2016) studied the open datasets, methods and metrics in telecom CCP. The authors found out that Support Vector Machines (SVM) out-performed Decision Tree and Neural Networks in churn classification during CCP. A comparison study carried out by also revealed SVM-POLY as the best classifier. However in Monani et al. (2016)'s survey still on CCP in telecom, Neural networks were found to outperform the rest of the techniques in terms of performance. The authors' research also recommended survival analysis as a solution for determining the possibility of a customer churning at a particular moment in time.

3. Materials and Methods

This study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). The present literature was subjected to selection criteria. For inclusion, the topic of research was about customer churn prediction models in the telecom industry. The selected studies were of quantitative research and published from 2014 to 2017. The language of publication was English. The criteria also ensured that selected studies exhibited transparency about data used in the analyses thus reliability in their statistical results. This selection criteria therefore excluded studies that did not meet any of the above stated requirements.

3.1 Search Strategy

The research strategy is used to ensure that literature is from three academically recognized databases. These included; Science direct, Emerald Insight, and Springer. For specificity, certain key words were used during the search for relevant literature. These were; "customer churn", "churn prediction" and "customer churn prediction". All the key words were always combined with "in the telecommunications industry," "telecom industry" or "in the mobile industry" to ensure the domain of interest. The initial search yielded 744 articles, from which only 30 were included in the final study. The flow of the screening process is shown in Figure 2.

3.2 Characteristics of Included Studies

This section looks at the features of the 30 selected studies. As shown in Figure 3, the highest number of studies was carried out in 2017 followed by 2016. Science direct produced the highest number of studies followed by Springer as shown in Figure 4.

Different churn prediction techniques were used in the selected studies. In some study cases, a combination of these was used purposely for model enhancement, accuracy comparison and also proof of superiority of proposed techniques. In a churn prediction model of prepaid customers, Azeem, Usman and Fong (2017) compared their proposed fuzzy classifier techniques with other non-fuzzy techniques. For model improvement, Idris, Iftikhar and Rehman (2017) used a combination of genetic programming and AdaBoost thus forming GP-Adaboost with advantages of both its components. Table 1 shows the frequency of the techniques that reappeared in different studies for either comparison or model development. In this study, four mostly used prediction techniques are briefly described. Different data preparation methods were also used to make data suitable for model fitting. These included; Feature selection, Normalization, Undersampling, Noise removal and these are described in the sections that follow. Weka mining was the mostly used data mining tool.

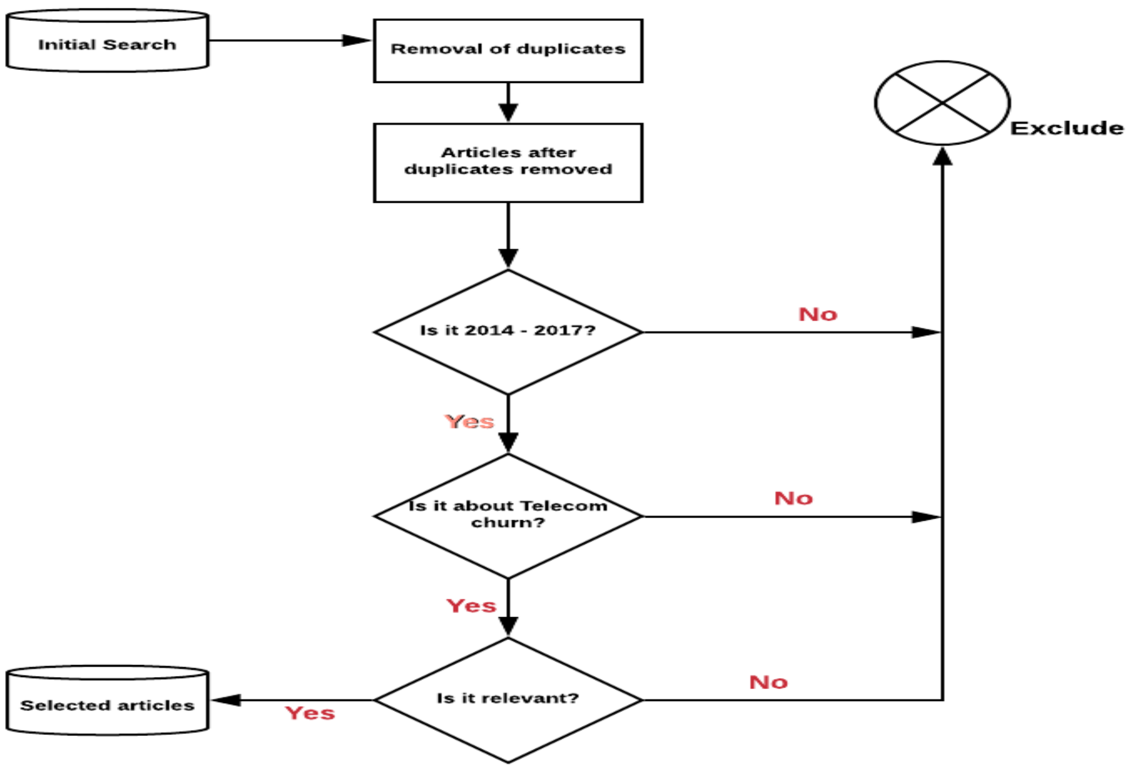

Figure 2: The Selection process and criteria.

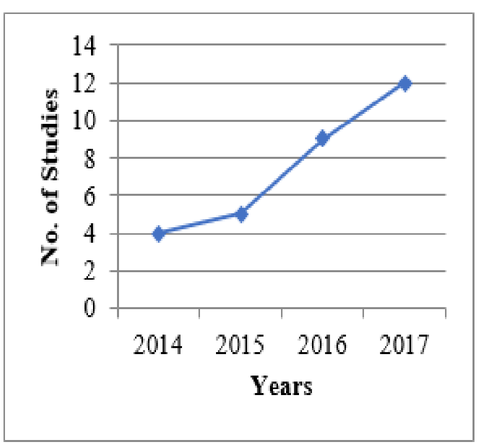

Figure 3: Studies carried out per year

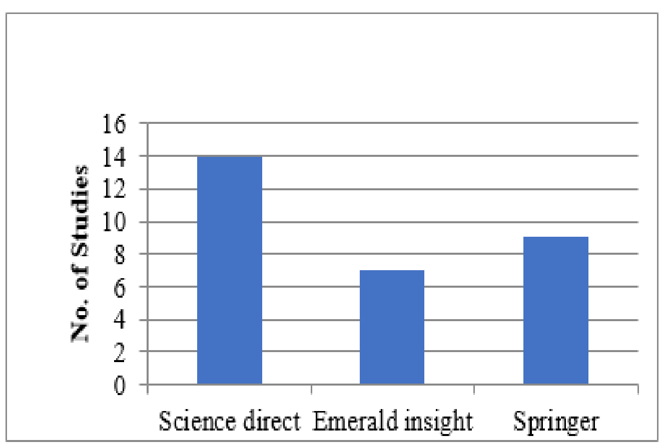

Figure 4: Contribution of studies by databases

4. Results and Discussion

This section looks at the findings of this study. The different data preparation techniques, churn prediction techniques, churn prediction challenges are the main contents discussed.

4.1 Data Preparation Techniques

Coussement, Lessmann and Verstraeten (2017) asserted that data preparation is aimed at turning data into a state suitable for further analysis. Different data preparation techniques were used in the selected studies. As shown in Figure 5, Feature Selection is the mostly used technique followed by Normalization and random Undersampling. This called for their discussion that follows in the subsections that follow. Noise removal is also very important in data preparation as confirmed by Azeem, Usman and Fong (2017), who used this technique to deal with missing values, outliers and unique characters that existed in telecom datasets which they used in their analysis.

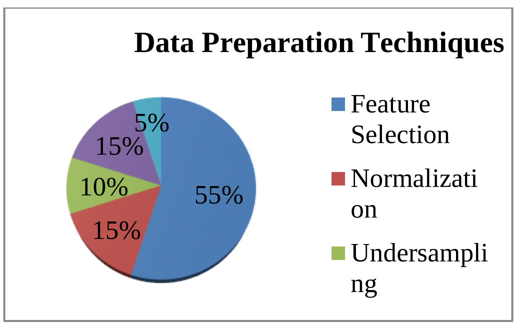

Figure 5: Data preparation techniques.

4.1.1 Feature Selection

In this method, Amin et al. (2017) stated that the most relevant attributes are selected from the many thus reducing computational costs. The choice of the relevant attributes is based on domain knowledge. As stated by Fei, Shuan and Yan (2017), attributes that are highly interdependent are screened so that those with little contribution to the dataset are removed. In churn prediction cases, attributes such as customer bills, call duration details and customer demographics are considered important attributes (Azeem et al., 2017). As a result, the model is left to work with those influential attributes which improves its predictive performance (Umayaparvathi & Iyakutti, 2016). The high dimensionality issue with telecom data is also solved by feature selection as acknowledged by Coussement, Lessmann and Verstraeten (2017) in their comparative study of data preparation techniques. In a survey about CCP methods, telecom datasets and metrics, Umayaparvathi and Iyakutti (2016) used feature selection to reduce the dimensionality of the data. Feature selection is therefore a good data preparation technique to employ especially when the size and dimensionality of data need to be reduced.

4.1.2 Normalization

This technique is used to deal with values having dynamic ranges by assigning them a small specified range. Zhao et al. (2017) used the min-max normalization method in which the scale assigned to values ranges from 0 to 1, thus making all the values lie in this range. In this method, the minimum and maximum values take on 0 and 1 respectively. This range makes it easier to visualize and analyse data. Furthermore, normalization makes it easy for prediction techniques to perform effectively.

4.1.3 Undersampling

Telecom datasets are characterized by class imbalance issues because the focus is on churners who are always the minority group (Idris et al., 2017). In order to balance the 7.3% minority (churners) class with the majority (non-churners) class, the above authors used Undersampling to reduce samples in the majority group instances by choosing distinctive ones. This technique therefore reduces the bias that naturally exists in the distribution of telecom datasets, thus reducing CCP model training bias. Amin et al. (2017) used Undersampling to balance churners with the non-churners in the training dataset. However, Gui (2017) disregarded Undersampling because a lot of meaningful information was lost with the application of this technique during his analysis of the telecom imbalanced data issue.

Table 1: Frequency of techniques usage.

Technique	Frequency
Support Vector Machines	7
Neural Networks	7
Decision Tree	5
Logistic Regression	4
Naïve Bayes	5
AdaBoost	3
K-Nearest Neighbour	3
Hybrid Firefly	2
K-Local Maximum Margin	2
Bagging	3
Fuzzy Classifiers	3
Particle Swarm Optimization	2
Total	**46**

4.2 Churn Prediction Techniques

A variety of churn prediction techniques were used as shown in Table 1. In certain studies, a combination of two or more techniques were used for model development and comparison purposes. Support vector machines emerged the mostly used technique, similar to Umayaparvathi and Iyakutti (2016)'s findings but contrary to Hashmi, Butt and Iqbal (2013)'s. The reason behind this rise in Support Vector Machines' usage is mainly due to their applicability in both regression and classification (Kumar & Chandrakala, 2017). Neural Networks have also been used with the same frequency as Support Vector Machines. This is due to their performance consistence even with large datasets. In this section, Support Vector Machines, Neural networks, Decision tree and Naïve Bayes are discussed considering the fact that they were used the most as evidenced by their frequency of usage in Table 1.

4.2.1 Support Vector Machines (SVM)

Khodabandehlou and Zivari Rahman (2017) described SVM as a machine learning technique for solving linear and non-linear classification problems. The authors asserted that SVM aims at minimizing the distance between hyper planes and classes to separate the classes as much as possible. In telecom CCP, SVM have been widely used and registered success (Yu et al., 2016). However, the technique was disregarded by Coussement, Lessmann and Verstraeten (2017) who claimed that it requires additional parameter tuning and often fails to give straightforward predictions. In this study, SVM registered the highest frequency of usage.

4.2.2 Neural Networks (NN)

Neural Networks work in a way that converts data into a brain neuron system (Monani et al., 2016), and are used in speech and pattern recognition as well as computer vision. The authors also noted its easy parallelization, continued learning even after application of the training set but to also observed that it is not easy to understand. In their survey on CCP techniques, the authors credited NN as better than most other methods. NN are also more capable for large datasets compared to other techniques. This justifies their increased adoption in the telecom churn prediction where large datasets are the order of the day. NN are however complex to comprehend and interpret.

As shown in Figure 6, the structure of a NN works on two major components, namely the processing elements and the weightage links connecting them. Neurons (also known as nodes or units) are the core processing elements which serve as functions. They receive information as input from other neurons. A neural network is divided into three basic layers with each containing neurons; an input layer, hidden layers and the output layer. The neurons in the input layer are the only ones that receive information from outside the neural network. This could be input data from any sources. The neurons in the hidden and output layers receive their input as output from other neurons. The neuron then multiplies this input with a weight and sums these multiplications. The sums are then passed to an activation function and the result becomes the input to another neuron the next layer. The process continues until the output layer produces the output as either yes or no (or 0/1) for the case of churn.

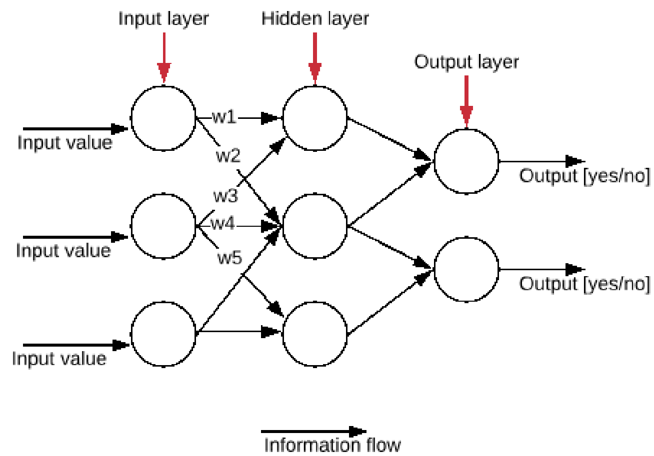

Figure 6: The structure of a neural network with one hidden layer. Each circle represents a neuron while the connecting arrows are the weightage links with weights (w1, w2,...).

4.2.3 Naïve Bayes (NB)

Fei, Shuan and Yan (2017) acknowledged this technique as a simple probabilistic method that is normally used in dichotomous datasets. It works on the conditional probability principle of the Bayesian rule (Babu, 2016), thereby using and analyzing all the variables individually even when they are independently important. NB classifies by treating all the variables independently in their influence to the churn variable. This has infact attracted various research works to modify NB to account for the correlations among the variables. Given a set of independent variables (X_1,……,X_m) and a dependent categorical (churn) variable (Y), then the probability for a certain output (yes/no for churn) given that the different variables have occurred and been measured will be;

$$P\left({}^{Y}\!/_{(X_1,\dots\dots X_m)}\right) = \frac{P\left({}^{(X_1,\dots.,X_m)}\!/_{Y}\right)P(Y)}{P(X_1,\dots.,Xm)}$$

4.2.4 Decision Trees (DT)

Existing either as categorical or regression trees, DT use a tree-like graph, considering all the possible outcomes from which a graphical representation of decision rules is drawn accordingly (Vafeiadis, Diamantaras, Sarigiannidis, & Chatzisavvas, 2015). As shown in Figure 7, class labels are represented by the tree leaf nodes while the conjunctions leading to those classes are represented by the tree branches. DT use if else statements and each internal node contains a test on a feature. The branches represent the result of a test while the topmost node represents the root node. Despite difficulty in its creation and proneness to over fitting, Monani et al. (2016) confirmed its simplicity to be understood and to generate rules in their survey on CCP in the telecom industry.

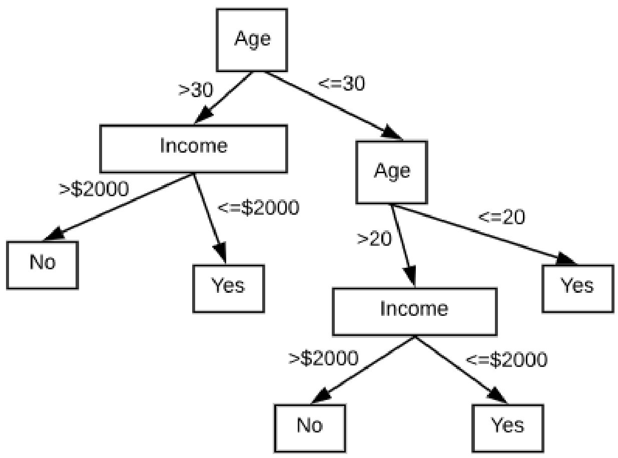

Figure 7: A simple DT structure with the leaf nodes having classes as either Yes or No.

4.3 Churn Prediction Challenges

Researchers in the selected studies highlighted some challenges encountered during the studies. This section looks at two challenges that cut across the selected studies.

In telecom churn prediction datasets, the class of interest is the churners group which is always the minority thus causing a data imbalance problem (Zhu, Baesens and vanden Broucke, 2017; Amin, Anwar, et al., 2017; Gui, 2017; Óskarsdóttir et al., 2017). In their empirical comparison among the different class imbalance solutions, Zhu, Baesens and vanden Broucke (2017) stated that this problem affects the performance of the prediction models because the bias caused by the majority class also affects the model performance. The studies selected in this review had datasets with churn rates ranging from 1.5% to 9% which adds to the class imbalance problem. The above authors also highlighted solutions to the class imbalance problem by categorizing them into data-based, algorithm-based and ensemble-based solutions. A data-based solution of random oversampling was also used by Azeem, Usman and Fong (2017) and Óskarsdóttir et al. (2017) to balance the datasets used in which the churn rates were 9% and 8.5% respectively. Oversampling is however not an ideal technique for handling class imbalance issues given the huge nature of telecom datasets, as Idris, Iftikhar and Rehman (2017) challenged it and proposed PSO based Undersampling. As the class imbalance issue continued to prevail, Amin, et al. (2017) handled it using the random Undersampling method.

Telecom datasets also come in the context of big data due to their large volumes, high velocity, veracity and variety as Li et al. (2015), Hafez (2016), Ahmed and Maheswari (2017) and Ahmed and Maheswari (2017) acknowledged. It therefore requires skilled personnel as well as complex big data tools and techniques to draw meaningful and actionable insights from it. Furthermore, the dimensionality of telecom data is high. In a review about the structural problems of big data, Storey and Song (2017) highlighted new technologies recently developed to handle big data such as Hadoop, Hbase and NoSQL, but also did not rule out the fact that these new technologies still have a low acceptance and adoption rate among big data users. This is because of the installation and training costs associated with these technologies even though they are open source platforms.

5. Research Findings

The research yielded findings in the field of telecom churn prediction with respect to the years considered.

- Support Vector Machines and Neural Networks are the mostly preferred prediction techniques. The reason for SVM high adoption is mainly due its applicability in both regression and classification problems while NN are preferred due to their abilities to perform well even with huge telecom data.
- Feature selection is the mostly used data preparation technique for dimensionality reduction
- Class imbalances and large datasets coupled with their high dimensions are the major challenges in telecom churn prediction
- Undersampling is preferred for handling data class imbalances

6. Conclusions

Customer churn in the telecom industry is indeed a serious issue that needs timely detection so that competitive retention strategies are taken. Telecom datasets should be handled carefully because of their imbalanced nature, large volumes and high dimensional structure. Feature selection is preferred for dimension reduction while Undersampling for class balancing. It is to the discretion of the company about which data mining tool to handle the large telecom datasets with. Different CCP methods have been proposed with Support Vector Machines and Neural Networks as the mostly preferred prediction techniques. However, ensembles of these techniques improve the prediction accuracy of the models because of the combined advantages of the components.

Further studies should focus on the profitability associated with attributes when dimension reduction is done. There is also a very big gap in real-time churn prediction for future studies to explore and this can be effectively achieved with the adoption the big data technologies. There is also need for more versatile approaches that consider the likely effects of national or regional economic changes on the behaviour of customers in ways that make these consider such changes. Natural language processing could also be used to detect churners using words and voice recognition to identify them based on these. Research is also needed to create proper visualizations of the huge telecom churn data. Visualizations speed up the process of discovering data patterns to reveal actionable insights.

References

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. 6(3). p.52–67.

Ahmed, A. A. Q., & Maheswari, D. (2017). Churn prediction on huge telecom data using hybrid firefly based classification Churn prediction on huge telecom data. Egyptian Informatics Journal. 18(3). p.215–220.

Akmal, M. (2017). Factors Causing Customer Churn: A Qualitative Explanation Of Customer Churns In Pakistan Telecom Industry.

Amin, A., Al-Obeidat, F., Shah, B., Tae, M. Al, Khan, C., Durrani, H. U. R., & Anwar, S. (2017). Just-in-time customer churn prediction in the telecommunication sector. Journal of Supercomputing. p.1–25.

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(December 2016). p.242–254.

Azeem, M., Usman, M., & Fong, A. C. M. (2017). A churn prediction model for prepaid customers in telecom using fuzzy classifiers. Telecommunication Systems. 66(4). p.603–614.

Babu, S. (2016). A Study on Efficiency of Decision Tree and Multi Layer Perceptron to Predict the Customer Churn in Telecommunication using WEKA. International Journal of Computer Applications. 140(4). p.26–30.

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. p.27–36.

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. p.1–13.

Federico, C. (2014). Data Preparation in the Big Data Era. Igarss 2014, (1), p.1–5.

Fei, T. Y., Shuan, L. H., & Yan, L. J. (2017). Prediction on Customer Churn in the Telecommunications Sector Using Discretization and Naïve Bayes Classifier. International Journal of Advances in Soft Computing and Its Applications. *9*(3).

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

Hafez, H. A. A. (2016). Mining Big Data in Telecommunications Industry : Challenges , Techniques , and Revenue Opportunity. *18*(1). p.4297–4304.

Hashmi, N., Butt, N. A., & Iqbal, M. (2013). Customer Churn Prediction in Telecommunication: A Decade Review and Classification. International Journal of Computer Science Issues. 10(5). p.271–282.

Idris, A., Iftikhar, A., & Rehman, Z. U. (2017). Intelligent churn prediction for telecom using GP-AdaBoost learning and PSO undersampling. Cluster Computing. p.1–15.

Khodabandehlou, S., & Zivari Rahman, M. (2017). Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior. Journal of Systems and Information Technology. 19(1/2). p.65–93.

Kumar, A. S., & Chandrakala, D. (2017). An Optimal Churn Prediction Model using Support Vector Machine with Adaboost. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2(1). p.225–230.

Li, H., Wu, D., Li, G. X., Ke, Y. H., Liu, W. J., Zheng, Y. H., & Lin, X. La. (2015). Enhancing Telco Service Quality with Big Data Enabled Churn Analysis: Infrastructure, Model, and Deployment. Journal of Computer Science and Technology. 30(6). p.1201–1214.

Mehwish, F. A., Zaffar, A. S., & Sumaira, J. M. (2017). Research Article Autonomous Toolkit to Forecast Customer Churn. International Journal of Current Research. Vol. 9. Issue 12. p.62999-63006.

Mitrović, S., Baesens, B., Lemahieu, W., & De Weerdt, J. (2017). On the operational efficiency of different feature types for telco Churn prediction. European Journal of Operational Research. 0. p.1–15.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Systematic Reviews and Meta-Analyses: The PRISMA Statement. Annulas of Internal Medicine. 151(4). p.264–269.

Monani, S., Gupta, S., Jain, S., & Mishra, V. (2016). Survey on Prediction of Customer Churn Analysis in a Telecommunications Industry. International Journal for Research in Engineering Application and Management, (07). p.22–26.

Olson, D. L., & Delen, D. (2008). Data Mining Process. Springer. p.10-35.

Óskarsdóttir, M., Bravo, C., Verbeke, W., Sarraute, C., Baesens, B., & Vanthienen, J. (2017). Social network analytics for churn prediction in telco: Model building, evaluation and network architecture. Expert Systems with Applications. 85. p.204–220.

Storey, V. C., & Song, I. Y. (2017). Big data technologies and Management: What conceptual modeling can do. Data and Knowledge Engineering. 108(February). p.50–67.

Umayaparvathi, V., & Iyakutti, K. (2016). A Survey on Customer Churn Prediction in Telecom Industry: Datasets, Methods and Metrics. International Research Journal of Engineering and Technology. p.1065–1070.

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. p.1–9.

Yu, R., An, X., Jin, B., Shi, J., Move, O. A., & Liu, Y. (2016). Particle classification optimization-based BP network for telecommunication customer churn prediction. Neural Computing and Applications. p.1–14.

Zhang, S., Zhang, C., & Yang, Q. (2010). Data preparation for data mining. Applied Artificial Intelligence. 17. 2003.

Zhao, L., Gao, Q., Dong, X. J., Dong, A., & Dong, X. (2017). K- local maximum margin feature extraction algorithm for churn prediction in telecom. Cluster Computing. 20(2). p.1401–1409.

Zhu, B., Baesens, B., & vanden Broucke, S. K. L. M. (2017). An empirical comparison of techniques for the class imbalance problem in churn prediction. Information Sciences. 408. p.84–99.