Journal of Applied Technology and Innovation (eISSN: 2600-7304)
vol. 1, no. 2, (2017), pp. 28-40

Literature Review of Data Mining Techniques in Customer Churn Prediction for Telecommunications Industry

Sherendeep Kaur
Faculty of Computing, Engineering & Technology
Asia Pacific University of Technology & Innovation
57000 Kuala Lumpur, Malaysia
Email: TP019638@mail.apu.edu.my

Abstract **-** Customer churn is one of the most critical issues faced by the telecommunications industry. In the telecommunications industry, it is more expensive to acquire a new customer as compared to retaining the current one. Hence, customer churn prediction is currently the main mechanism employed by the industry in order to prevent customers from churning. The objective of churn prediction is to identify customers that are going to leave the telecommunications service provider in advance. Customer churn prediction would allow the telecommunications service provider to plan their customer retention strategy. The high volume of data generated by the industry, with the help of data mining techniques implementation, becomes the main asset for predicting customer churn. Due to this reason, recent literature of different data mining techniques and most popular data mining algorithms for customer churn prediction are reviewed in this paper. Additionally, recent literature on newly developed algorithms based on the popular algorithms are also reviewed.

Index Terms **–** Data Mining, Big data analytics, Churn prediction

1. Introduction

The telecommunications industry in today's world is dealing with a major deficit towards their generated revenue due to the aggressive market competition (Umayaparvathi & Iyakutti, 2012). The telecommunications industry as any other service providers' industries consider customers to be the most crucial resource for them. However, the aggressive competition, such as lucrative retention polices offered by telecommunications service providers to attract new customers, has resulted major loss of customers in this industry as customers tend to leave one telecommunications service provider to another for this reason or another (Adwan et al., 2014). It is possible for telecommunications companies to focus more on acquiring new customers, however customer acquisition usually costs more compared to customer retention, and this will eventually lead to lower revenue (Chen, Fan & Sun, 2012). Hence, due to this reason, the telecommunications industry has shifted their focus to customer retention (Adwan et al., 2014).

The act of customers leaving one service provider to another service provider is called customer churn (Umayaparvathi & Iyakutti, 2012). Recently, customer churn prediction has become a highly-discussed domain. Many studies have analysed customer churn problem from many different viewpoints to design and recommend the best solution to the telecommunications industry (Shaaban et al., 2012).

Data mining techniques have been widely used as the solution to predict customer churn, by identifying the factors that are most likely contributing into customer churn in order to enable telecommunications service providers to take immediate action to prevent churning (Umayaparvathi & Iyakutti, 2012). The large volume of data, such as demographic data, billing information, call details, network details, among others help to ensure the accuracy of the data mining technique's application in the telecommunications industry (Umayaparvathi & Iyakutti, 2012). Among many data mining techniques for customer churn prediction, supervised data mining techniques are the most extensively explored. Supervised data mining techniques are appropriate when models to be developed can learn from labelled training data. Supervised data mining techniques consist of varied algorithms such as linear regression, neural networks, decision trees, k-nearest neighbours, genetic algorithms, Naïve Bayes, support vector machines (SVM) and others. (Shaaban et al., 2012)

In section 2, customer churn prediction is explained in detail and the objective of customer churn prediction as well as the types of churners are also discussed. In section 3, data mining is defined and the relation between data mining and knowledge discovery in database (KDD) is explained. Additionally, the steps required in KDD process are also explained in this section. In section 4, various data mining techniques and algorithms for customer churn prediction in telecommunications industry is presented. Lastly, the conclusion is presented in section 5.

2. Customer Churn Prediction

The objective of customer churn prediction is to predict the impending churners based on the predefined forecast horizon, assuming the data related with each subscriber in the network. According to Umayaparvathi & Iyakutti (2012), the customer churn prediction problem is normally characterized into three major stages, namely, training stage, test stage and prediction stage. In the training phase, the contribution for customer churn problem is from the historical data such as call details and personal and/or business customers' data, which has been obtained and retained by the telecommunications service providers. Furthermore, in the training stage, the labels are structured in the list of churners' records. In the test stage, the trained model with highest accuracy is tested to predict the churners' records from the actual dataset which does not contain any churn label. Lastly, in the prediction stage, which is also known as the knowledge discovery process, the problem is classified as predictive modelling or predictive mining.

Customer churn prediction helps the customer relationship management (CRM) to avoid customers who are expected to churn in future by proposing retention policies and offering better incentives or packages to attract the potential churners in order to retain them. Hence, the possible loss of the company's revenue can be prevented. (Umayaparvathi & Iyakutti, 2012)

Shaaban et al. (2012) stated that there are two types of churners, namely, involuntary and voluntary. Involuntary churners are the list of customers that are removed by the telecommunications service provider, itself, due to non-payment status, deception and non-usage of phone. Meanwhile, voluntary churners are the customers that decide to terminate their service with the respective telecommunications service provider. Involuntary churners are easy to be recognized; however, the voluntary churners are more difficult to be identified. Generally, the customer churn problem in the telecommunications industry is voluntary.

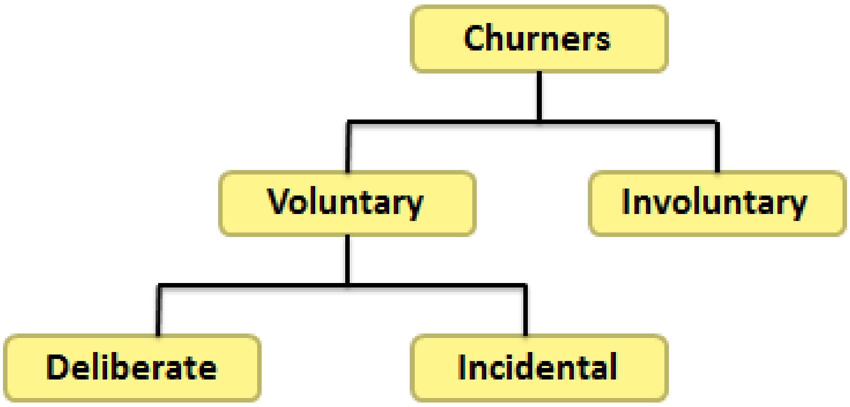

Figure 1: Types of Churners (Shaaban et al., 2012)

It can be seen in Figure 1 that voluntary churn is separated into two sub-categories which are deliberate churn and incidental churn. Deliberate churn is resulted from factors such as economic factors (example: price sensitivity), technology factors (example: more innovative technology is offered by another telecommunications service provider), poor customer service factors and others inconvenience related factors. Incidental churn is not caused by customers' plan, but because of sudden changes occurred in the customers' lives, such as changes in financial situation, geographical or relocation changes and others. (Shaaban et al., 2012)

3. Data Mining

Data mining is described as the process of determining and extracting valuable information in large databases (Verbeke, 2012). Data mining is an essential element of knowledge discovery in databases (KDD) process. KDD is a process which explains the steps that must be taken to ensure a thorough data analysis. KDD process contains five steps as illustrated in Figure 2.

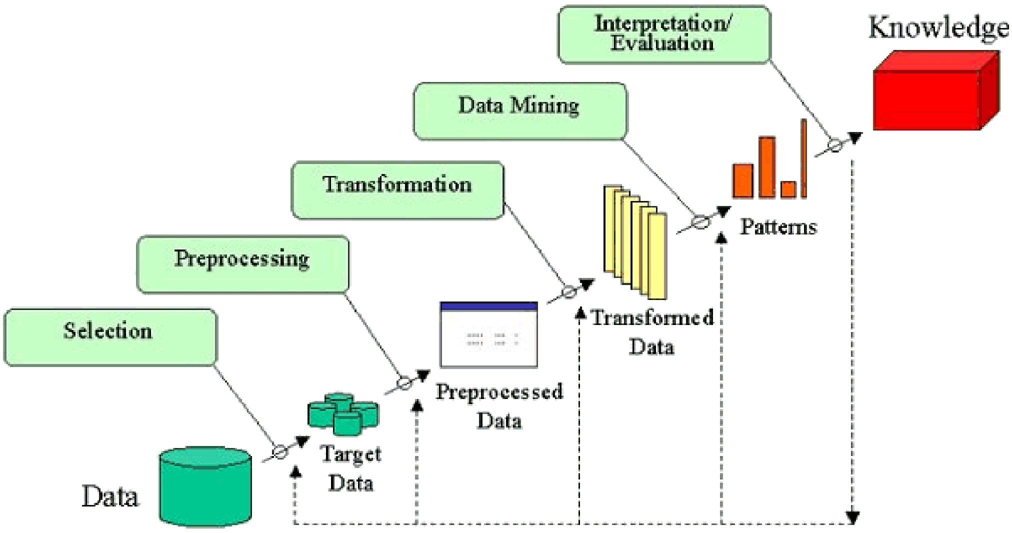

Figure 2: Steps of KDD Process (Tomar & Agarwal, 2014)

The first step involves the selection of the right data from numerous heterogenous data sources to create the target dataset. The right selection of attributes and variables is very important at this step in order to produce an accurate analysis. In data preprocessing step, missing values, outliers and inconsistency are removed and/or cleaned. This step is required to enhance the quality of the target dataset as well as the mining results. The transformation step focuses on transforming the variables from one format to another to ease the implementation of the data mining algorithms. The transformation step involves changing variables from one format to another to perform data mining. The application of data mining step has to be aligned with the objective of the entire KDD process, in this case, customer churn prediction. The last step is the interpretation and evaluation of the mined data, this step determines the success level of the purposed model and documents the knowledge discovery for future references. (Tomar & Agarwal, 2014). To effectively manage the churn prediction problem, different researches have implemented different data mining algorithms which is extensively reviewed in the following section.

4. Data Mining Techniques and Algorithms for Customer Churn Prediction

Azeem, Usman & Fong (2017) used fuzzy algorithm to predict customer churn in telecommunications industry. Recently, the fuzzy algorithm has gained popularity in telecommunications industry to conduct customer churn prediction as normally the data required for churn prediction is very noisy and fuzzy algorithm is very effective in managing noisy data. The dataset for this research was obtained from a telecommunications company based in South Asia. SVM, linear regression, C4.5, gradient boosting, Artificial Neural Network (ANN), random forest, AdaBoost and neural network were used for performance comparison with fuzzy algorithm. The fussy algorithm utilized numerical variables and domain knowledge in the variable selection stage. The previous researches conducted on customer churn prediction, mainly focused only on post-paid customers. However, with the implementation of fuzzy algorithm, call details record variable of prepaid customers was also considered in this research and the effectiveness of predicting prepaid customer churn was highly measured. The performance evaluation results showed that the fuzzy algorithm was more accurate and effective as compared to other algorithms with Area Under the Curve (AUC) rate of 98% and True Positive score of 0.68.

Ismail et al. (2015) proposed the Multilayer Perceptron (MLP) algorithm to predict customer churn. The dataset for this research was obtained from one of the telecommunications company in Malaysia. The MLP algorithm was compared with two statistical algorithms which are logistic regression and multiple regression. The performance of the algorithms was measured based on sensitivity, accuracy and specificity. The variables used for this analysis were customer demographic variables, customer relationship variables, billing variables and usage variables. MINITAB software was utilized for constructing the multiple regression and logistic regression algorithm and MATLAB software was utilized for constructing the MLP neural network algorithm. The performance evaluation result showed MLP neural network as the best algorithm for customer churn prediction with an accuracy of 91.28%. Meanwhile, multiple regression produced accuracy prediction of 78.84% and logistic regression produced prediction accuracy of 75.19%. Additionally, the sensitivity showed 93.5% for MLP neural network, 83.98% for multiple regression and 71.80% for logistic regression. Furthermore, the specificity of showed 88.28% for MLP neural network, 71.90% for multiple regression and 79.66% for logistic regression. The overall finding result suggested that MLP neural network algorithm is the most suitable approach for customer churn prediction as compared to conventional statistical algorithms.

Vafeiadis et al. (2015) conducted customer churn prediction by comparing the commonly used data mining algorithms such as ANN, decision trees, SVM, Naïve Bayes and logistic regression. These algorithms were then compared with their enhanced versions in order to boost their performance further. The main objective of this study was to evaluate the suitability of the state of data mining algorithms for the customer churn problem. The telecommunications dataset for this research was obtained from UCI Machine Learning Repository. The dataset was transformed to MLC++ format in order to enable the performance comparison of the algorithms and their enhanced versions. The evaluation was performed using Monte Carlo simulation at various settings of each algorithm. The enhanced algorithm methodology used for this study was AdaBoost.M1. The objective of this methodology is to boost the performance of the algorithms' performance through a combination of weak algorithms, such as decision tree and ANN, in order to obtain better decision. Monte Carlo simulation for this dataset revealed that the combination of ANN and decision tree algorithms as well as SVM achieved the highest accuracy compare to others. The enhanced algorithm methodology is proven to be very useful for telecommunications industry as it can improve the rate of accuracy for customer churn problems, especially for large customers' data.

Whilst most researches on customer churn prediction focused more on data mining algorithms than identifying the most important variables to be used for these algorithms, Adwan et al. (2014) conducted a research on customer churn prediction by applying a MLP algorithm. MLP neural network not only predicts customer churn, but also provides insights on the significance of each input variable used for customer churn prediction. The MLP neural network has an extensive range of implementation for classification and prediction problems in business and industrial domains. This research used real data which was provided by Umniah, a major telecommunications network in Jordan. Confusion matrix was used to assess the MLP neural network algorithm. Additionally, k-cross validation with k=5, was implemented in order to provide a better understanding on how well the algorithms would perform when new data is presented. Furthermore, to understand the importance of each variable towards the churn, change of error (COE) and weights contribution on the network were implemented. COE grades input variables, fitting to the variation of some quality measures when each input is erased from the dataset in the training process. Weight contribution in the network is applied in order to obtain comparative frequencies for the variables. There were three variables that appear on both COE and weight contribution in the network, namely, total monthly fees, total of international outgoing calls and 3G services. These findings can be valuable to the customer relationship management department in the telecommunications industry as it allows them to plan more effective strategies to manage customer churn. Nevertheless, there is a noticeably disadvantage of the COE method in which the noise and inconsistency in the training dataset can minimize its consistency.

Keramati et al. (2014) used a dataset obtained from an Iranian mobile company for their research. They compared the performance of four algorithms which are decision tree, ANN, k-nearest neighbours and support vector machine. After analysing the algorithms' performance and understanding their special structures, a new algorithm was proposed in order to improve the evaluation metric value. The proposed algorithm, which was named as "Best Hybrid Methodology" in this study, showed above 95% accuracy for recall and precision were easily attained. Additionally, compared to the four algorithms, the new proposed algorithm also has better tuning parameter that can be easily deployed to predict as per the analyst's requirement. If an analyst needs to analyse which customers are likely to churn and which customers are likely to have the least tendency to churn, the algorithm can be tuned to produce the desired analysis. Furthermore, this study also introduced a new dimensionality reduction techniques in order to extract the most important set of variables to be involved in data mining process.

Kim, Jun & Lee (2014) conducted customer churn prediction analysis by presenting an innovative variable, called the network variable, which was acquired from network analysis. The network variable was calculated through the propagation process. The spreading activation was employed in the propagation process. Additionally, the network variable as well other customers' personal details variables were used as the input. Unlike other researches on customer churn prediction, this research actually measured the numerous features of all existing churners who influence potential churners. The dataset was acquired from a telecommunications company, which involved the call detail records data and customers' personal detail data. Logistic regression and ANN were employed in this research. To accomplish better performance evaluation of the algorithms, two hypotheses were applied. The first hypothesis was that existing churners have an influence on the potential churners in the same group. The second hypothesis was that churn date and centrality influence the existing churners in leading potential churners. Group identification was employed to the network variable and the propagation process was conducted in the same group to investigate the first hypothesis. The result showed that the prediction performance of both algorithms after group detection was still similar to non-group detection result, but the propagation process time was faster. The second hypothesis performance was evaluated by adjusting the churners' existing energy considering the centrality and churn date. The result showed that adjusting the churners' existing energy enhanced the prediction, and both algorithms also performed better when the existing energy was adjusted.

Abbasimehr, Setak & Soroor (2013) proposed two stage structure for customer churn prediction. The first stage is the identification stage and the second stage is the data mining algorithm stage. In this research, social network based variables were included along with conventional variables from customer relationship management database. The neuro fuzzy algorithm was used in the data mining process. The effectiveness of the two types of neuro fuzzy algorithms which are locally linear neuro-fuzzy (LLNF) with locally linear model tree (LoLiMoT) and the adaptive neuro-fuzzy inference system (ANFIS) were explored. Additionally, the two neuro fuzzy algorithms were compared with two neural network algorithms, namely, MLP and radial basic function (RBF). The dataset used for this research was obtained from the Teradata Centre at Duke University. The customers' variable in the dataset was clustered by using k-means algorithm and the customers' data that were on the top cluster was used in the data mining algorithm stage. The result showed that neuro fuzzy algorithms performed better compared to the neural network algorithms.

Brandusoiu & Toderean (2013) performed a customer churn prediction analysis by utilizing dataset obtained from Department of Information and Computer Science of University of California. IBM SPSS (Statistical Product and Service Solutions) was used as the technology to mine the data. There were no missing values identified from the dataset and there was a perfect correlation (R=1) between some variables SVM algorithm was used as a data mining algorithm for this research. As a requirement of SVM algorithm, The "yes-es" in the dataset was cloned by balancing and by boosting the training set in order to have an equal distribution with the "no-es". The SVM algorithm was trained by using four kernel functions which are RBF, Linear, Polynomial (POL) and Sigomid (SIG). The result showed that POL performed the best out of the four kernels with overall accuracy of 88.56%. However, overall the four kernels predicted around 80% percent accuracy. Based on the predictors that have higher significance in scoring the kernel performance, the customer relationship management can plan different marketing approach to retain potential churners. In spite of this, the study was only conducted with only one algorithm, comparison between few algorithms would reveal different results and higher accuracy rate could have been achieved.

Kirui et al. (2013) added a new subcategory of variables in order to enhance the accuracy of customer churn prediction in the telecommunications industry. The new subcategory of variables which are call pattern description variables, call pattern changes description variables, contract-related variables, were originated from the statistical traffic data and customers profile data. The dataset used for this study was obtained from European telecommunications company. Bayesian Network and Naïve Bayes were used to assess the performance of predictive significance of the added new variables. The evaluation results were then compared using C4.5 decision tree algorithm. True churn rate and false churn rate were achieved from the algorithms implementation. Bayesian Network and C4.5 decision tree showed a strong distinction of significance of every added subset of variables. However, Naïve Bayes showed that most of the added subset of variables performed almost evenly. In spite of this, C4.5 decision tree showed better accuracy performance on the dataset used. Yet, according to (Kirui, 2013), the false churn rate and the true churn rate are better measures as compared to accuracy rate for the customer churn prediction case. Nevertheless, this study did not address the class imbalance problem of the initial datasets. Hence, the minority class instances were not clearly identified even though they might actually achieve high overall accuracy.

Qureshi et al. (2013) tested logistic regression, decision tree (including CHI-squared Automatic Interaction Detector (CHAID), Exhaustive CHAID, Classification and Regression Trees (CART) and Quick, Unbiased and Efficient Statistical Tree (QUEST)), ANN and k-mean clustering on the data set acquired from an online source (http://www.customer-dna.com/). Additionally, re-sampling method was used in this research to tackle a very general problem in telecommunications industry which is class imbalance problem. Recall, precision and F-measures were used to evaluate the performance of different prediction algorithms. Furthermore, to determine the important variables as predictors, p-value below 0.05 and Pearson correlation were used. The re-sampled dataset gave unbiased result compare to the dataset with class imbalance. After implementation of all the algorithms, the results showed that Exhaustive CHAID, a variant of decision trees, was the most accurate for the dataset with accuracy of 70%. To further boost the accuracy, five new variables derived from some of the existing variables, were introduced to the dataset. The overall accuracy after including the new variables was indeed increased to 75.4%.

Chen, Fan & Sun (2012) proposed a data mining algorithm named the Hierarchical Multiple Kernel Support Vector Machine (H-MK-SVM) for both statistical and longitudinal behavioural data. The telecommunications dataset used for this study was from a mobile services company provided by the Centre for Customer Relationship Management of Duke University. To train the H-MK-SVM, a three-stage algorithm were established and applied. The longitudinal behavioural data for this study was not transformed in the training process. It was used directly as an input for the data mining algorithm without any clustering step as commonly done in standard contexts. Additionally, the training procedure of the H-MK-SVM was also a variables selection procedure because the spare non-zero coefficient correlation to the selected variables. The H-MK-SVM algorithm constructed a classification formula by calculating the coefficients of both longitudinal and statistical data. To compare the H-MK-SVM algorithm's performance, ten other algorithms, namely, the MK-SVM, SVM, least squares SVM (LS-SVM), decision tree, logistic regression, feed-forward ANN, RBF Neural Network, random forest, AdaBoost, and the proportional hazard model (Cox), were used in this study. The performance result utilizing the Lift and the AUC measures revealed that the H-MK-SVM performed effectively on both imbalance as well as balance class data compare to other algorithms. A collaborative data mining algorithm such as H-MK-SVM is a developing framework especially for customer churn prediction in the telecommunications industry as it performs effectively with large volume of data.

Huang, Kechadi & Buckley (2012) focused on deriving new sets of variables from the initial variables to predict customer churn. The new sets of variables which are aggregated call details, account information, Henley segmentation, bill information, dial types, payment information, service information, line information, complaint information, among others, were compared to the existing variables. The dataset itself, was obtained from a telecommunications company in Ireland. The prediction was conducted by using seven different algorithms which are linear classifiers, decision trees, logistic regression, ANN, SVM, Data Mining by Evolutionary Learning (DMEL) and Naïve Bayes on both derived set of variables and the existing set of variables. The result revealed that the new proposed set of variables were more effective for the churn prediction as compared to the existing set of variables. Furthermore, the result showed that decision tree and SVM with a low ratio were compatible for predicting the true churn rate and the false churn rate. Additionally, DMEL algorithm was revealed to be the weakest algorithm for customer churn prediction as it was not compatible for large dataset with high dimension. However, this study did not include enough derived variables in the data mining process which could improve the prediction accuracy.

Phua et al. (2012) predicted customer churn as well as customers win-backs in the near future based on datasets obtained from a telecommunications company. The customers involved were individuals and small medium enterprises (SMEs). The prediction objectives for this study was not only on identifying the likely churners with good accuracy but also identifying these churners within a short time period of consequent three months. An appropriate computational strategy was used in order to find fixed patterns to predict churners and possible win-back. Additionally, to attain dependencies for the stronger original variables, a few derived variables were created. The class imbalance correction was also performed by using under-sampling and over-sampling strategies in order to enhance the accuracy for minority class instances. Tree classifiers algorithms were used in this study, namely, ADTree, decision stump, RepTree, J48, TreeLMT, random forest, bagging + decision stump, bagging + simple cart, simple cart. Additionally, Naïve Bayes and classification via regression algorithms were also used as comparison. The performance evaluation result of the algorithms showed that random forest and simple cart performed the best with the highest accuracy prediction as compared to other algorithms.

Shaaban et al. (2012) proposed a model which consists of six stages which are identify problem domain, data selection, investigate data set, classification, clustering and knowledge usage to perform data mining for customer churn prediction, as shown in Figure 3; in which, the classification phase created two types of customers, namely, churners and non-churners. The clustering phase created three clusters which are utilized for evaluation of retention strategy for further research. The clustering step is not limited to three clusters only, the number of clusters is subject to the type of knowledge usage. The knowledge usage obtained the clusters in order to provide solution to retain each type of churners. Churners can be grouped according to many criteria based on customer dissatisfaction and/or profitability. The dataset used for this study was obtained from an anonymous mobile service provider. The dataset was separated into a training set (80%) and a test set (20%). decision tree, ANN, SVM were used in this research. The confusion matrix showed that ANN and SVM predicted 83.7% of accuracy while decision tree predicted 77.9% of accuracy for this dataset. (Shaaban et al., 2012).

Umayaparvathi & Iyakutti (2012) conducted churn prediction analysis with dataset obtained from a data mining competition, PAKDD-2006. The dataset was aggregated into training data and test data for six months' period, as prediction models need historical data of customer behaviour for a specific period of time in order to predict the customer behaviour in the future. Most relevant variables for algorithm implementation were selected. The variables later were categorized into 4 major groups which are customer demography, customer care service, bill and payment and call detail record. These variables are trained in order to implement the decision tree algorithm and ANN algorithm. The performance of the algorithms was tested based on the counts of test data. A confusion matrix was established for both algorithms based on demographic variables in order to find the predictive accuracy and error rate of the algorithms. The confusion matrix revealed that decision tree algorithm surpassed the ANN algorithms for customer churn prediction, as the error rate was lower by 0.4 % and the accuracy was higher by

0.4% as well compare to ANN algorithms based on the datasets used. However, this study did not combine other variable groups in the testing process. The result could have been different and could also predict better accuracy if other variable groups were tested as well.

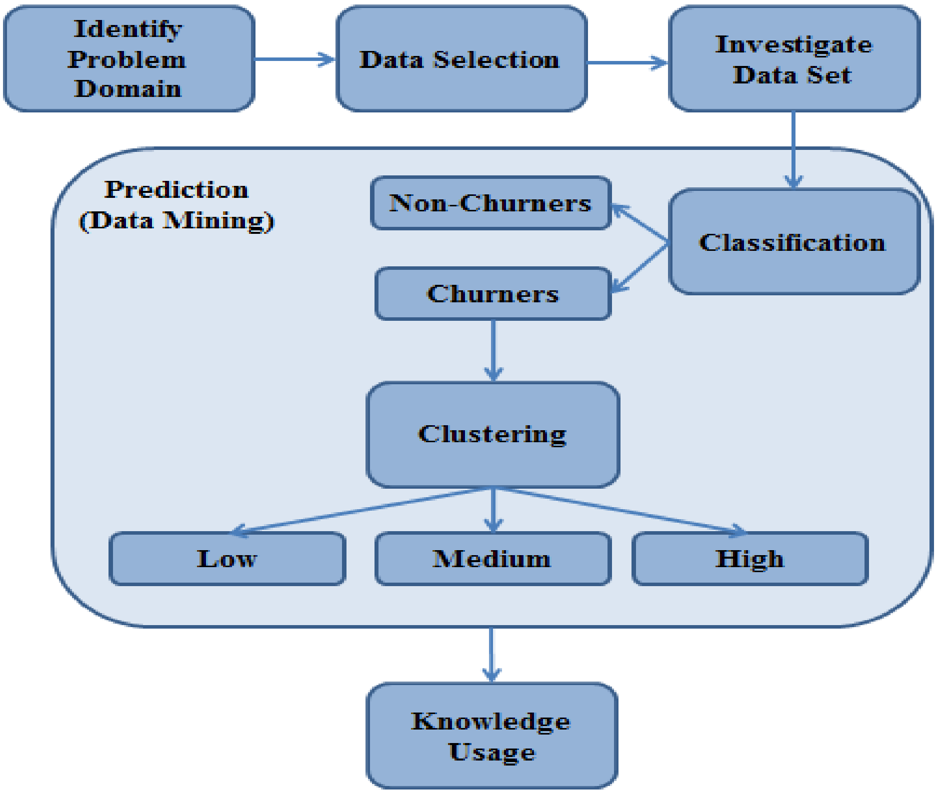

Figure 3: Proposed Model for Customer Churn Prediction (Shaaban et al., 2012)

5. Conclusion

Customer churn prediction is one of the most crucial missions in telecommunications industry. The aggressive market of telecommunications industry has forced the service providers to employ the best data mining algorithms which produce most accurate prediction in order for them to stay competitive in the market.

Many data mining algorithms have been reviewed and SVM, Bayes Network, decision tree, ANN, amongst others, were found to be the most popular algorithm in customer churn prediction. Some researchers also have combined few algorithms and established an innovative algorithm in order to produce better accuracy rate. Additionally, the enhanced algorithm methodology such as AdaBoost was found to be very valuable, as it can enhance the accuracy performance of weak algorithms. However, the accuracy performance of each algorithm differs in every research. This is due to the different dataset used and the different input variables chosen for the experiment.

Most of the literature focused more on data mining algorithms, but only a few of them focus on distinguishing the important input variables for churn prediction to be used for data mining algorithms implementation. Additionally, only noticeably one literature that had actually combined social network based variables in the input variables for data mining algorithms implementation. Moreover, the class imbalance problem was found to be not addressed on some of the literatures.

In future, this literature review can be used as a foundation for customer churn prediction analysis. Few of the most prominent algorithms can be tested on new dataset. Additionally, external environment factors such as social factors can be considered as one of the input variables for churn prediction as it is proved to have an impact on the accuracy result.

References

Abbasimehr, H., Setak, M. & Soroor, J. (2012) A Framework for Identification of High-Value Customers by Including Social Network Based Variables for Churn Prediction Using Neuro-Fuzzy Techniques. International Journal of Production Research. 51(4). p.1279-1294.

Adwan, O., Faris, H., Jaradat, K., Harfoushi, O. & Ghatasheh, N. (2014) Predicting Customer Churn in Telecom Industry Using Multilayer Preceptron Neural Networks: Modeling and Analysis. *Life Science Journal*. 11(3). p.1-7.

Azeem, M., Usman, M. & Fong, A.C.M. (2017) A Churn Prediction Model for Prepaid Customers in Telecom Using Fuzzy Classifiers. *Telecommunications Systems*. p.1-12.

Brandusoiu, I. & Toderean, G. (2013) Churn Prediction in the Telecommunications Sector Using Support Vector Machines. *Annals of the Oradea University Fascicle of Management and Technological Engineering*. 22(1). p.19-22.

Chen, Z.Y., Fan, Z.P. & Sun, M. (2012) A Hierarchical Multiple Kernel Support Vector Machine for Customer Churn Prediction Using Longitudinal Behavioral Data. *European Journal of Operational Research*. 223(2). p.461-472.

Huang, B., Kechadi, M.T. & Buckley, B. (2012) Customer Churn Prediction in Telecommunications. *Expert Systems with Applications*. 39(1). p.1414-1425.

Ismail, M.R., Awang, M.K., Rahman, M.N.A. & Makhtar, M. (2015) A Multi-Layer Perceptron Approach for Customer Churn Prediction. *International Journal of Multimedia and Ubiquitous Engineering*. 10(7). p.213-222.

Keramati, A., Marandi, R.J., Aliannejadi, M., Ahmadian, I., Mozzafari, M. & Abbasi, U. (2014) Improved Churn Prediction in Telecommunications Industry Using Datamining Techniques. *Applied Soft Computing*. 24. p.994-1012.

Kim, K., Jun, C.H. & Lee, J. (2014) Improved Churn Prediction in Telecommunications Industry by Analyzing a Large Network. *Expert Systems with Applications*. 41(15). p.6575-6584.

Kirui, C., Hong, L., Cheruiyot,W. & Kirui, H. (2013) Predicting Customer Churn in Mobile Telephony Industry Using Probabilistic Classifiers in Data Mining. *International Journal of Computer Science Issues*. 10(2). p.165-172.

Phua, C., Cao, H., Gomes, J.B. and Nguyen, M.N. (2012) Predicting Near-Future Churners and Win-Backs in the Telecommunications Industry. arXiv preprint arXiv:1210.6891.

Qureshi, S.A., Rehman, A.S., Qamar, A.M., Kamal, A. & Rehman, A. (2013) *Telecommunications Subscribers' Churn Prediction Model Using Machine Learning*. In Digital Information Management (ICDIM), 2013 Eighth International Conference on. 10th September 2013. IEEE. p. 131-136.

Shaaban, E., Helmy, Y., Khedr, A. & Nasr. M. (2012) A Proposed Churn Prediction Model. *International Journal of Engineering Research and Applications*. 2(4). p.693-697.

Tomar, D. & Agarwal, S. (2014) A Survey in Pre-Processing and Post-Processing Techniques in Data Mining. *International Journal of Database Theory and Application*. 7(4). p. 99-128.

Umayaparvathi, V. & Iyakutti, K. (2012) Applications of Data Mining Techniques in Telecom Churn Prediction. *International Journal of Computer Applications.* 42(20). p.1-5.

Vafeiadis, T., Diamantaras, K.I., Sarigiannidis, G. & Chatzisavvas, K.C. (2015) A Comparison of Machine Learning Techniques for Customer Churn Prediction. *Stimulation Modelling Practice and Theory*. 55. p.1-9.

Verbeke, W., Dejaeger, K., Martens, D., Hur, J. & Baesens, B., (2012) New Insights into Churn Prediction in the Telecommunications Sector: A Profit Driven Data Mining Approach. *European Journal of Operational Research*. 218. p.211-229.