UDC 004.65:654
Survey Paper

Review of Data Mining Techniques for Churn Prediction in Telecom

Vishal Mahajan *vmahajan@hcl.com*
HCL, Technologies,
Noida, India

Dr. Richa Misra *richa.misra@jaipuria.ac.in*
Jaipuria Institute of Management,
Noida, India

Dr. Renuka Mahajan *rmahajan@amity.edu*
Amity University Uttar Pradesh,
Noida, India

Abstract

Telecommunication sector generates a huge amount of data due to increasing number of subscribers, rapidly renewable technologies; data based applications and other value added service. This data can be usefully mined for churn analysis and prediction. Significant research had been undertaken by researchers worldwide to understand the data mining practices that can be used for predicting customer churn. This paper provides a review of around 100 recent journal articles starting from year 2000 to present the various data mining techniques used in multiple customer based churn models. It then summarizes the existing telecom literature by highlighting the sample size used, churn variables employed and the findings of different DM techniques. Finally, we list the most popular techniques for churn prediction in telecom as decision trees, regression analysis and clustering, thereby providing a roadmap to new researchers to build upon novel churn management models.
Keywords: Customer Churn, Telecom, Churn Management, Data Mining, Churn Prediction, Customer retention

1. Introduction

Today's customers have a seemingly endless supply of information at their fingertips. Smartphones, for example, enable much faster access to brand, product and price-comparison information. As a result, companies in multiple industries are having difficulty in attracting and retaining customers. Due to the rapid technological advances and increased competition, customers have multiple options to choose and this has become a challenge for telecom operators. Companies are losing a lot of revenue due to switching by their existing customers. This process is called "Churn".

Churn in telecom industry, means measurement of customers that change service or service provider over a given period of time [13], [16], [26], [102], [21], [79],[11],[60]. Churn can be both voluntary and involuntary. Voluntary churn happens when existing customer leaves the service provider and joins another service provider, while in involuntary churn customer is asked by the service provider to leave due to reasons like non-payments etc. [34],[57]. Voluntary churn can be sub-divided into: incidental churn and deliberate churn [35]. Incidental churn occurs, not because the customers planned for it but because something happened in their lives e.g. a change in financial condition, change in location etc. Deliberate churn occurs for reasons of technology (customers wanting a newer or better technology, price sensitivity, service quality factors, social or psychological factors and convenience reasons) [69].

The term 'Churn Management' is a process, used by telecom companies to retain their profitable subscribers. Similarly, [53] explains churn management in the telecom industry as the procedure of retaining the most important customers for the company. He also emphasized to predict how each customer will react to specific offers and predict which customers will be positively influenced. Authors of [93] stated that if an existing subscriber terminates a contract with one service provider and becomes a subscriber of another service provider, then this subscriber is called as 'lost' customer or 'Churn' customer.

Data mining predicts future trends and behaviors, which helps businesses to become more proactive and allows them to take knowledge driven decisions. Data mining can answer business questions that traditionally were too time-consuming to resolve. Data mining (DM) is the science of analyzing large databases to find patterns and trends [61], [37], [88], [29]. It is defined as "the nontrivial extraction of valid, novel, potentially useful and understandable information from data" [37], [99], [36].

The effective research cannot be accomplished without critically studying what already exists in the form of general literature and specific studies pertaining to churn.

In this study, we tried to address the following questions:

- What are the specific customer-based application areas to which DM methods have been applied?
- What are the commonly used DM methods applied in these domains?
- Over what kind of churn dimensions and sample size, do the methods operate in telecom sector?
- To summarize the most popular DM techniques used in churn prediction models in telecom sector.
- To indicate fertile areas for further research work in the field.

Such a study helps the researcher to avoid overlapping efforts and make new basis for novice researchers. The following sections provide the review of various literatures on understanding the variety of data mining strategies used for building churn prediction models. This paper aims at reviewing the research intensity during the year of 2000 to October, 2014. The research methodology is presented in section 2. The review of various mining techniques is stated in section 3 and section 4 gives the discussions. The conclusions and future scope of the paper are shown in section 5.

2. Research Methodology

In the field of data mining, there's extensive literature that is spread across several domains. We started the literature survey in September 2014. Consequently, to capture as many citations as possible relevant to customer churn prediction models, the following online journal databases were searched with keywords 'Churn Prediction model' and 'Data Mining'.

- Emerald
- Kluwer and Wiley
- Science Direct
- IEEE Transaction
- Elsevier
- SCOPUS
- Springerlink
- IEEE Xplore
- EBSCO (electronic journal service)

The electronic search was supplemented by manually searching journals, periodicals, abstracts, indexes, directories, research reports, conference papers, market reports, annual reports and books on marketing, customer relationship management, data mining and knowledge discovery. This extraction resulted in 1826 citations. Out of these, 1511 citations were excluded as irrelevant. Only those articles that had been published in business intelligence, knowledge discovery or data mining related journals were selected, as these were the most appropriate for data mining research and the focus of this review paper. The full papers of the remaining 315 citations were then evaluated to select those primary studies that were published starting from year 2000 till date. These criteria excluded 209 studies and left 106 in the review. They originated from thirteen countries, published in various languages between 2000 and 2014. Of these studies, unpublished working papers were excluded and finally 100 relevant to the purpose of review aided in recognizing how previous researchers' choice on data mining techniques affects their research discoveries. Because journals are considered as the most reliable source of research [75], so firstly, some renowned online journal databases were explored to get a comprehensive academic literature on the topic. Here is a list:

List of prominent Research Journals and Reports (International and Indian)

- Abhigyan
- Advanced Data Mining and Applications
- Advances in Knowledge Discovery and Data Mining
- Bell Labs Technical Journal
- Cellular Operators association of India Reports & statistics
- Cellular operators association of India, Trends and Development.
- Data Mining and Knowledge Discovery
- Asia Pacific Financial Markets
- European Journal of Marketing
- European Journal of Operational Research
- Expert Systems with Applications
- IEEE International Conference on Data Mining
- IEEE Transactions on Evolutionary Computation
- IEEE Transactions on Knowledge and Data Engineering
- Indian Journal of Marketing.
- International conference on Computational Science and Its Applications
- International conference on Extending database technology: Advances in database technology
- Knowledge-Based Systems
- Marketing Intelligence and Planning
- National Telecom Policy 2012. Retrieved from Regulatory Authority of India.
- Telecom Regulatory Authority of India –Statistics & performance
- The Journal of Mobile Communication, Computation and Information
- The Mckinsey Quarterly
- Omega The International Journal of Management Science
- US mobile - Telecommunications Policy

3. Usage of Data Mining in predicting churn in Telecommunication

The purpose of data mining (DM) is to analyze large set of data to retrieve meaningful information. Customer churn prediction has been raised as a key issue in many fields such as telecommunication [44], [84], [96], [46] , [41],[56], Credit Card [75], Internet Service Providers [43], [55] Electronic Commerce [17] , [57], [30], [63], [94] Retail Marketing [19], Newspaper publishing companies[28], [8], banking [62],[22],[102],[4] and financial services [65].

One of the important applications of data mining is Churn Analysis in telecom industry. It is used to predict the behavior of customers who are most likely to quit the services of existing provider and join new service provider. Understanding the current and past trends, behavior and planning for the future is important in business. Hence, data mining applications play an important role in decision making and providing prediction on future estimates.

Technically, data mining is the process of finding correlations or patterns among fields in large databases. Key data mining functionalities can be classified as follows: multivariate statistical analysis (regression analysis), relationship mining (frequent pattern mining algorithms), clustering, classification (decision trees, neural networks), prediction and outlier detection [10], [38].

Predictive modeling is essentially concerned with foreseeing how the customer will behave in the future by investigating their past behavior [40]. Anticipating customers who are likely to churn is one example of the predictive modeling. It is used in analyzing Customer Relationship Management (CRM) data and data mining (DM) to deliver customer-based models that depict the probability that a customer will take a specific action [73]. These actions could be sales, marketing and customer churn/retention related. There are many models that can used to distinguish between churners and non-churners in an organization.

Researcher [74] classified Customer Relationship Management dimensions into four sets i.e. Customer Identification, Customer Attraction, Customer Retention and Customer Development using popular data mining functions such as Association, Classification, Clustering, Forecasting, Regression, Sequence Discovery and Visualization.

According to [49], neural networks have wide range of applications for prediction and classification problems in industrial and business domains. Authors [77] used neural networks techniques to predict the customer churn. Authors used the randomly selected 5000 customers from a Jordian telecommunication company. They utilized customers billing information (monthly fee, call rate, SMS fee), usage behavior (minutes of usage, number of SMS), users past churning status and plan type (3G). They found that monthly fees, total minutes of usage and 3G services have been the most influencing factors to predict the churn.

As explained by [67], logistics regressions, classification, clustering and decision tree are very successful for predicting the customer churn. Author used survival analysis and hazard function to investigate the customers who are highly likely to churn and the time when they will churn. Survival analysis provides the probability of survival of a customer after an observation period, while hazard function is used to predict that customer will churn during a time period. He used a sample data set of 41,374 from a telecom company in one of the state in USA and used customer demographic data (age, gender, income etc.), customer internal data (plan type, billing agency, billing disputes, number of weekly calls, national and international call billing etc.) and customer contract records for their study. Author opined that using these techniques they could find 90% of the churners.

Authors [44] conducted a study to predict the churner in the Taiwan telecom industry, which had become a major focus of the industry. They considered customer demography (age, tenure, gender), billing and payment information (billing amount, monthly fee, overdue payment), call details (call duration, call type), customer care services of the customers from one of the telecom company including the churners for a period of one year. They segmented the customers into various clusters using K-means clustering based on amount, tenure, outbound call usage, inbound call usage, and payment rate. Authors found that the corporate users has high probability to churn, may be due to change in job. They suggested that users who do not make call to other users on the same network have a high probability to churn. They also found that users whose contract is going to expire in near future have more probability to churn.

Author [93] conducted a research in Turkey, where telecom sector was suffering huge customer loss, to find out what types of customers are switching the service provider and what are the reasons behind that. Authors analyzed records of 1000 customers for a period of 6 months and used Logistic Regression and Decision Trees mining techniques. They used the subscriber's usage as parameter to predict the churn and found that if subscriber does not have any discount package, then there is 75% likely hood that subscriber will churn. Authors also found that other important factor responsible for churn is number of incoming calls and long distance calls. Subscribers who receive maximum calls from subscribers using same service provider are less likely to churn than subscribers using other service provider.

Authors in [55] used the clustering technique to predict the churn. They used the demographic, billing and usage data like frequency of usage, minutes of usage and volume of data usage pattern of the subscribers from an internet operator in Tehran for their work. They selected subscribers registered with in a particular month and then collected their information for a time span of 8 months. They used the k-means clustering technique to create the clusters based on frequency of usage, minutes of usage and volume of data usage and found that billing and usage features has the highest effect on churn prediction on the churn while demographical information has least effect on churn prediction.

In [54], authors used decision tree data mining technique to predict the customer churn for Malaysian telecom service provide. They used length of service, area and total of more than 10 minutes of customer engagement parameter for their investigation. Authors discovered that rural and urban users have different churning behavior. They found that if the subscriber belongs to sub-urban area and is engaged by customer services less than 10 minutes, they have high probability of churning than rural subscriber.

Authors [47] used classification data mining techniques to predict the churn behavior of subscribers from Satara, Maharashtra, India with a focus on post-paid subscribers only. They selected a sample of 895 users from various categories like business, private and government. Authors used call related information such as number of calls made, duration of calls, different number called for 2 weeks as parameters for their research. Authors stated that calling pattern for non-churning subscribers remains same for the period while for churning subscriber's numbers of calls in first few days are less and then it increases significantly.

Authors in [100] used classification techniques to find the factors affecting the churn and calculated the profits generated from retention. They used data from Europe, North America and East Asia for their study, collected over a period of three to six months. Authors suggested that customer churn prediction is more applicable in post-paid subscribers as lots of information is available about these users rather than pre-paid, where most of the subscribers are unknown. Authors found that small sample variables helps to predict churning more accurately and oversampling does not improve the churn prediction performance. They explained that since numbers of churner are far less than non-churners, this poses a problem for classification techniques to create powerful class distribution. They considered top 10% of the customers with highest probability to switch for their investigation. They inferred that any campaign for retaining customers will be profitable only when it is targeted towards small fraction of top customers and saves lots of money for the company. Authors also found that most of previous researches used classification techniques on a single data set for predicting churn and classification technique to be used for churn prediction is still an open research problem.

Authors [89] employed neural network based approach to predict the customer churn and found that neural networks of medium size their performance is better, when different neural network's topologies were investigated. Authors used dataset of 2427 customers from repository of Machine Learning Databases at the University of California for their investigation. They used state, tenure, area code, plan type, number of voice messages, number of calls in day, number of minutes used in a day, daily total charge, number of calls in evening, total evening call charges, total night calls, total night charges, number of international calls, total international call charges and number of calls to customer service variables for their research.

Authors [92] utilized Markov Logic Networks techniques to find the effect of word of mouth on churning of subscribers and their switching. They used the sample data set from a telecom provider for 2,645 customers including their call detail records for a period of 8 months. The sample includes the customer who has churned as well. Authors used various customer attributes including usage of data services, number of calls, type of mobile set, contract type, tenure, usage trend, plan changes, service center calls, customer demographics, including age and gender. Authors found a strong relationship between word of mouth and the churn behavior of the customer.

Authors [85] applied a social network based approach to predict the churn of subscribers. They divided the subscribers into various clusters based on social groups. They investigated the interactions between the members within a cluster, to know the status of each member. They provided a churn score to each member based on the churn score of the group he belongs to. They used statistical model to provide a churn score to each social group. They used only call data for their study and found that social leader can significantly impact the churn with in their groups. They also found that leader has 3 times greater probability of churn compared to other members of the group.

Authors [31] conducted a study to find out the factors which encourages customers to churn in Iran. Authors found that usage of service, customer satisfaction and demographic attributes has impact on subscriber churn. For their investigation, they used the data of 3150 subscribers from an Iranian service provider. The authors used Local Linear Model to predict subscriber churn. They used the level of customer dissatisfaction, customer demographic attributes, level of use and cost of churn for subscriber.

Authors [60] used probabilistic data mining Naiye Bayes and Bayesian network and decision tree technique to predict the customer churn. For their study, they collected the data of the subscribers from a European telecommunication provider for a period of three months. Authors used customer profile, traffic details, contract-related features (tenure), call patterns features (number of calls, call duration), and calls pattern changes features (change in frequency of use, change in minutes of use, change in activity). They used random sampling technique to select the sample for study. They reported that traffic details, call pattern features and change in call patterns play an important role in predicting customer churn than user profile data. They also found that probabilistic data mining classifier Naiye Bayes and Bayesian network have more accuracy in predicting the churn compared to decision tree.

Authors [50] used J48 and C5.0 decision trees as classification mining techniques to predict the potential churners, so that companies can use better retention strategies. For their research authors used data of 3333 including some churned customers from a telecom company in one of the cities in South India. Authors used 10 customer parameters for their research purpose, namely account number, area code, voice mail service, number of minutes per day, number of calls per day, daily call spend, International call duration, International number of calls and churn. Authors compared the accuracy of the two techniques on the same set of data and found that C5.0 classification technique is more efficient and accurate than J48 decision tree technique.

Authors [32] employed hybrid model using Logistic Regression in parallel with Voted Perceptron for classification, and combined with clustering for predicting churn in mobile subscribers. Authors used a dataset of 2000 subscribers from an Asian telecom provider for their investigation. For their research work authors used billing amount, location, price, tenure and age parameters for investigation. They found that subscriber churn has direct relationship with higher usage and low tenure and effect negatively. Also they found that monthly billing also impact the subscriber churn, higher the billing, higher the probability of churn.

Authors [2] conducted a survey on various data mining techniques being used in telecom industry for predicting the subscriber's churn. Authors analyzed Neural Networks, Decision Tree and regression techniques (Linear regression, Logistic regression, Naive Bayes Classifier and K-nearest neighbor's algorithm). Author found that Decision Tree based techniques are more accurate than regression based techniques. They also found that Neural network based mining approach can give better results compared to decision tree and regression based mining techniques provided the data size and attributes are carefully selected.

4. Discussions

Table 1 shows various DM techniques (decision trees, neural networks, clustering, association analysis, support vector machines, clustering and others) that are used for predicting customer churn from 2000 and 2014 in different domains e.g. banking, newspaper, retail and credit risk analysis.

Author (Year)	References	Technique Used	Industry
Ahn at al. (2006)	[1]	logistic regressions	Telecom
Almana, Aksoy & Alzahrani(2014)	[2]	neural networks, decision tree and regression	Telecom
Antreas (2000)	[3]	confirmatory factor analysis , clustering	Banking
Au et al. (2003)	[5]	Genetic algorithms	Telecom
Ballings & Poel (2012)	[6]	classification and logical regression techniques	Newspaper
Benjamin et al. (2012)	[7]	discriminant & multivariate analysis	Telecom
Buckinx & Poel (2005)	[13]	logistic regressions and neural networks	Retail
Burez & Poel (2009)	[15]	logistic regression and markov chains random forests	Television
Burez et al (2007)	[14]	Markov chain, Logistic regression	Pay TV company
Chen & Ching (2007)	[18]	regression	Telecom
Chiang et al. (2003)	[23]	association rules	Banking
Chueh (2011)	[25]	fuzzy correlation analysis	Telecom
Coussement & Poel (2008)	[26]	support vector machines, random forests logistic regression	Newspaper
Datta et al. (2001)	[27]	decision tree	Telecom
Fasanghari & Keramati (2011)	[31]	local linear model	Telecom
Georges & Shuqin (2014)	[32]	logistic regression, clustering & classification	Telecom
Huang et al (2010)	[42]	neural network, decision tree	Wireless telecom
Hung et al. (2006)	[44]	classification (decision tree, neural network) clustering (k-means)	Telecom
Hwang H., Jung and Suh (2004)	[45]	Logistic regression, decision tree, neural network	Wireless telecom
Kamalraj & Malathi (2013)	[50]	decision tree and classification	Telecom
Kavipriya & Rengarajan(2012)	[51]	discriminant analysis, multiple regression	Telecom
Kim & Yoon (2004)	[57]	logistic regression	Telecom
Kirui et al. (2013)	[60]	decision tree	Telecom
Lariviere & Poel (2004)	[64]	hazard model survival analysis	Banking
Mallikarjuna, Mohan & Kumar (2011)	[68]	discriminant analysis	Telecom
Morik and Kopck (2004)	[70]	Decision tree, support vector machines	Insurance
Mues et al. (2004)	[72]	decision diagrams	Credit Risk Evaluation
Piotr (2008)	[80]	Rough-sets	Telecom
Poku, Zakari & Sonali (2013)	[81]	regression	Hotel
Rajkumar & Rajkumar (2010)	[83]	factor analysis	Telecom
Richter, Tov &	[85]	clustering	Telecom

Author (Year)	References	Technique Used	Industry
Slonim (2011)			
Sathish et al.(2011)	[87]	clustering	Telecom
Tamaddoni et al (2009)	[91]	neural networks, decision tree	Telecom
Torsten, Martin & Krishnan (2011)	[92]	Markov logic networks	Telecom
Verbeke (2011)	[96]	C4.5, ant miner, support vector machines and logistic regression	Telecom
Verbeke (2012)	[95]	classification	Telecom
Wei & Chiu (2002)	[98]	classification (decision tree)	Telecom
Wei & Chiu (2002)	[98]	decision tree	Telecom
Xia & Jin (2008)	[101]	support vector machine	telecom
Zhu et al (2009)	[103]	Bayesian networks, support vector machines	Wireless telecom

Table 1: Decade overview of Industry wise Research Techniques used

According to [39], following are the popular techniques have been reviewed in the light of academic literature

- Decision trees: These are most popular prediction models [71]. They are the trees like formations that represent sets of choices. These choices create 'if-then-else' rules for classifying the dataset [24], [9], [19] and [58].
- Regression Analysis: Regression analysis is next popular technique and is used for the investigation of relationships between variables. Regression analysis is done in order to evaluate the influence of some explanatory variable on the dependent variable. [78], [82], [97], [6].
- Neural Network: Neural network is a mathematical model, which is based on biological neural networks, which processes information using a connectionist approach to computation. [90], [99], [86].
- Cluster analysis: It attempts to discover natural groupings of observations in the data [48], [36].

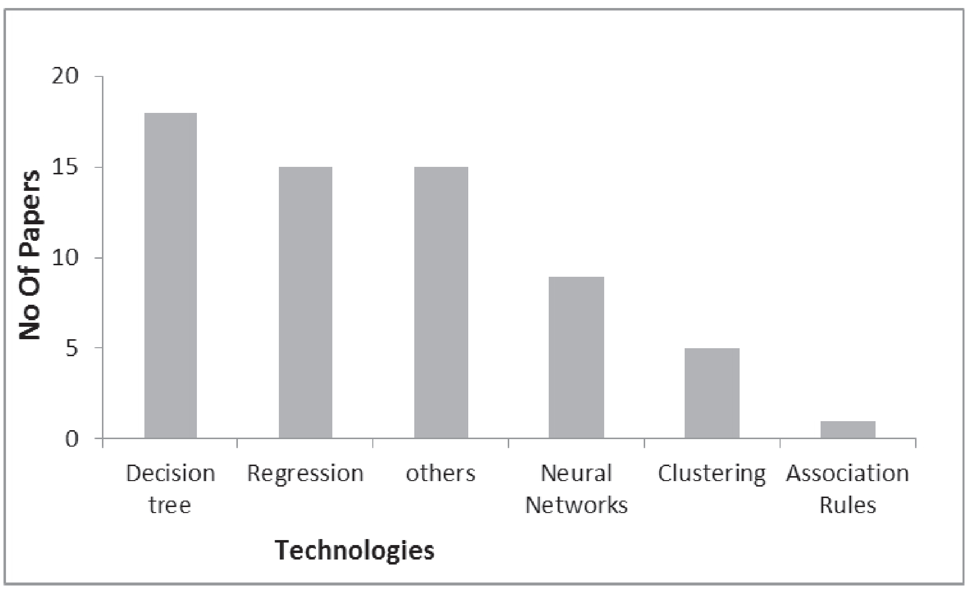

Figure 1: Technology wise Research Papers in Customer Churn Prediction

5. Conclusions and Future Scope

The findings of the survey show various DM techniques that are used for predicting customer churn from 2000 and 2014 in different domains e.g. banking, newspaper (media), retail and credit risk analysis. It outlines current trends, dimensions used and challenges of DM applications in telecommunications, for the researchers who are beginning to review this field. Based on literature review, various mining techniques used for churn so far, are classified in Table 1 above. The most popular DM techniques are decision tree, regression, neural network and clustering as depicted in Figure 1 above. However, there is no clear common consensus on the prediction technique to be used on the data collected [100]. Further, due to the cost involved, most of the existing studies involved in survey are using small data sample of customer records [52], [12], [33], [66], [59], which may undermine the reliability and validity of analysis results. It means an empirical study with a significant larger data set with added dimensions may increase the reliability of result.

It is suggested that the future development of DM techniques can become more problem-oriented and specific to 'churner type' prediction required. Moreover, the hybrid models can be introduced and compared with existing models. It will help in designing multidimensional customer dataset and devising new churn management techniques specific to different datasets and different geographical locations. The decision making based on analysis from DM techniques can make churn prediction even more accurate and will provide valuable insights to the cellular industry technique.

References

[1] Ahn, J.H.; Han, S.P.; Lee, Y.S. Customer churn analysis*:* Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecomm Policy*. 30, pages 552-568, 2006

[2] Almana, A.M.; Aksoy, M.S.; Alzahrani, R. A Survey on Data Mining Techniques in Customer Churn Analysis for Telecom Industry. *International Journal of Engineering Research and Applications*. 45, pages 165-171, 2014

[3] Antreas, D. Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research*. 47(3), pages 191–207, 2000.

[4] Au, W. H.; Chan K. C. C. Mining fuzzy association rules in a bank-account database, *IEEE Transactions on Fuzzy Systems*, 11, pages 238–248, 2003

[5] Au, W. H.; Chan, K. C. C.; Yao, X. A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction, *IEEE transactions on evolutionary computation,* 7(6), pages 532-545, 2003.

[6] Ballings, M.; Poel, D.V. D. Customer event history for churn prediction. How long is long enough? *Expert Systems with Applications*, 39(18), pages 13517–13522, 2012.

[7] Benjamin, O.; Mesike, G.; Bakarea, R.; Omoera, C.; Adeleke, I. Discriminant Analysis of Factors Affecting Telecoms Customer Churn. *International Journal of Business Administration*, 32, pages 59-67, 2012.

[8] Benoit, D. F.;Coussement, K.; Poel, D. V. D. Improved marketing decision making in a customer churn prediction context using generalized additive models, *Expert Systems with Applications*, 37(3), pages 2132–214, 2010.

[9] Berry, M. J. A.; Linoff, G. S. *Data mining techniques second edition – for marketing, sales, and customer relationship management*, 2004.

[10] Berson, A.; Smith, S.; Thearling, K. *Building data mining applications for CRM*. New York, NY. McGraw-Hill, 2000.

[11] Bhambri, V. Data mining as a Tool to Predict Churn Behaviour of Customers, GE-*International Journal of Management Research IJMR*, 2(3), pages 59-69, 2013.

[12] Bolton, R.; Kannan, P. K.; Matthew, B. Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value, *Journal of the Academy of Marketing Science*, 28(1), pages 95–108, 2000.

[13] Buckinx, W.; Poel, D. V. D. Predicting online-purchasing behavior, *European Journal of Operational Research*, 166(2), pages 557–575, 2005.

[14] Burez J.; Poel, D. V.D. CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services", *Expert Systems with Applications*, 32(2), pages 277–288, 2007.

[15] Burez, J.; Poel, D. V. D. Handling class imbalance in customer churn prediction, *Expert Systems with Applications*, 36(3), and pages 4626-4636, 2009.

[16] Chandar, M.; Laha, A.; Krishna, P. Modeling churn behavior of bank customers using predictive data mining techniques, *National conference on soft computing techniques for engineering applications*, pages 24-26, 2006.

[17] Changchien, S. W.; Lee, C. F.; Hsu, Y. J. On-line personalized sales promotion in electronic commerce, *Expert Systems with Applications*, 27(1), pages 35–52, 2004.

[18] Chen, J.S.; Ching, R.K.H. The effects of mobile customer relationship management on customer loyalty: Brand image does matter. 40th Annual Hawaii *International Conference on System Sciences HICSS'07, Hawaii*, 405, pages 2502-2511, 2007.

[19] Chen, M. C.; Chiu, A. L.; Chang, H. H. Mining changes in customer behavior in retail marketing, *Expert Systems with Applications*, 28(4), pages 773–78, 2005.

[20] Chen, Y. L.; Hsu, C. L.; Chou, S. C. Constructing a multi-valued and multi labeled decision tree, *Expert Systems with Applications*, 25(2), pages 199–209, 2003.

[21] Chen, Z.-Y.; Fan, Z.P.; Sun, M. A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data, *European Journal of Operational Research*, 223(2), pages 461–472, 2012.

[22] Chiang, D. A.; Wang, Y. F.; Lee, S. L.; Lin, C. J. Goal-oriented sequential pattern for network banking churn analysis, *Expert Systems with Applications*, 25, pages 293–302, 2002.

[23] Chiang, W.; Lee, L. Goal-oriented sequential pattern for network banking churn analysis", *Expert Systems with App*lications, 25(3), pages 293–302, 2003.

[24] Chu, B. H.; Tsai, M. S. ; Ho, C. S. Towards a hybrid data mining model for customer retention, *Knowledge-Based Systems*, 20(8), pages 703–718, 2007.

[25] Chueh, H. Analysis of marketing data to extract key factors of telecom churn management. *African Journal of Business Management*, 5(20), pages 8242-8247, 2011.

[26] Coussement, K.; Poel, D.V. D. Churn prediction in subscription services. An application of support vector machines while comparing two parameter-selection techniques, *Expert Systems with Applications*, 34(1), pages 313–327, 2008.

[27] Datta, P.; Masand, B.; Mani, D.; Li, B. *Automated Cellular Modeling and Prediction on a Large Scale*, Issues on the application of data mining,14(6), pages 485- 502, 2001.

[28] Douglas, S.; Agarwal, D.; Alonso, T. Mining customer care dialogs for daily news. *IEEE Transactions on Speech and Audio Processing*, 13(5), pages 652-660, 2005.

[29] Dunham, M. H. *Data Mining: Introductory and Advanced Topics*. Prentice Hall PTR Upageser Saddle River, NJ, USA ©2002, ISBN: 0130888923, 2006.

[30] Etzion, O.; Fisher, A.; Wasserkrug, S. E-CLV A modeling approach for customer lifetime evaluation in e-commerce domains, with an application and case study for online auction, *Information Systems Frontiers*, 7, pages 421–434, 2005.

[31] Fasanghari, M.; Keramati, A. Customer Churn Prediction Using Local Linear Model Tree for Iranian Telecommunication Companies. *Journal of Industrial Engineering, University of Tehran*, Special Issue, pages 25-37, 2011.

[32] Georges, D.; Shuqin, C. A Hybrid Churn Prediction Model in Mobile Telecommunication Industry, 4(1), pages 55-62, 2014

[33] Gerpott, T.J.; Rams, W.; Schindler, A. Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. *Telecommun Policy*, 25(4), pages 249-269, 2001.

[34] Gordon, S. Survival Data Mining for Customer Insight. Retrieved from. *http.//www.data-miners.com/resources*, 2004.

[35] Gotovac, S. Modeling Data Mining Applications for Prediction of Prepaid Churn in Telecommunication Services, 51(3), pages 275-283, 2010.

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

[37] Han, J. ; Kamber, M. *Data Mining, Concepts and Techniques*, 2nd ed, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2006..

[38] Han, J.; Kamber, M. *Data Mining, Concepts and Techniques*, Academic Press, San Diego, 2001.

[39] Hashmi, N.; Butt, N. A.; Iqbal, M. Customer Churn Prediction in Telecommunication: A Decade Review and Classification, *International Journal of Computer Science Issues*, 10, pages 271-282, 2013.

[40] Hassouna, M. *Agent Based Modelling and Simulation*: An Examination of Customer Retention in the UK Mobile Market. PhD thesis, Brunel University, UK, 2012.

[41] Huang, B.; Kechadi, M. T. ; Buckley B. Customers churn prediction in telecommunications, *Expert Systems with Applications*, 39(1), pages 1414–1425, 2012.

[42] Huang, B.Q.; Kechadi, T.M.; Buckley, B.; Kiernan, G.; Keogh, E.; Rashid, T. A new feature set with new window techniques for customer churn prediction in land-line telecommunications, *Expert Systems with Applications*, 37(5), pages 3657–3665, 2010.

[43] Huang, P.; Lurie, N.H.; Mitra, S. Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods, *Journal of Marketing*, 73(2), March: 55 – 69, 2009.

[44] Hung, S-Y; Yen, D. C.; Wang, H-Y. Applying data mining to telecom churn management, *Expert System with Applications*, 31(3), pages 515–524, 2006.

[45] Hwang, H.; Jung, T.; Suh, E. LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, *Expert Systems with Applications*, 26(2), pages 181–188, 2004.

[46] Idris, A.; Rizwan, M.; Khan, A. Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection

strategies, *Journal of Computers and Electrical Engineering*, 38(6), pages 1808–1819, 2012.

[47] Jadhav, R.; Pawar, U. Churn Prediction in Telecommunication Using Data Mining Technology. I*nternational Journal of Advanced Computer Science and Applications - IJACSA* 2, pages 17-19, 2011.

[48] Jain, A. K.; Murty M. N.; Flynn P. J. Data clustering: A review. *ACM Computing Surveys*, 31(3), pages 264–323, 1999.

[49] Joseph, P. B. Data Mining with Neural Networks. *Solving Business Problems from Application Development to Decision Support*. McGraw-Hill, Inc., Hightstown, NJ, USA, 1996.

[50] Kamalraj, N.; Malathi, A. Applying Data Mining Techniques in Telecom Churn Prediction. *International Journal of Advanced Research in Computer Science and Software Engineering*, 310, pages 363-370, 2013.

[51] Kavipriya, T.; Rengarajan, P. User's Level of Satisfaction with Mobile Phone Service Providers - With Special Reference to Tiruppur District, Tamil Nadu. *National Monthly Refereed Journal of Research in Commerce; Management*, 19, pages 35-52, 2012.

[52] Keaveney, M. Customer switching behavior in service industries. *An exploratory study. Journal of Marketing*, 59(2), 71-82. doi. 10.2307/1252074 , 1995

[53] Kentrias, S. Customer relationship management. *The SAS perspective*, www.cm2day.com, downloaded: 2001

[54] Khalida, O.; Sunarti, M.; Norazrina, H; Faizin, B. Data Mining in Churn Analysis Model for Telecommunication Industry. *Journal of Statistical Modeling and Analytics*, 11 pages 19-27, 2010.

[55] Khan, A. A.; Jamwal, S. ; Sepehri, M.M., Applying Data Mining to Customer Churn Prediction in an Internet Service Provider, *International Journal of Computer Applications*, 9(7), pages 8-14, 2010

[56] Kim, B. An empirical investigation of mobile data service continuance: Incorporating the theory of planned behavior into the expectation–confirmation model. *Expert Systems with Applications*, 37(10), pages 7033-7039, 2010.

[57] Kim, H. S.; Yoon, C. H. Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market, *Telecommunications Policy*, 28(9–10), pages 751–765, 2004.

[58] Kim, J. K.; Song, H. S.; Kim, T. S.; Kim, H. K. Detecting the change of customer behavior based on decision tree analysis, *Expert System with Applications*, 22(4), pages 193–205, 2005.

[59] Kim M.K.; Park M.C.; Jeong D.H. *The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services*, 28(2), pages 145–159, 2004

[60] Kirui, C.; Hong, L.; Cheruiyot, W.; Kirui, H. Predicting Customer Churn in Mobile Telephony Industry Using Probabilistic Classifiers in Data Mining, *International Journal of Computer Science Issues IJCSI*, 10(2), No. 1, pages 165-172, 2013.

[61] Klosgen, W.; Zytkow, J. *Handbook of data mining and knowledge discovery*, New York. Oxford University Press, 2002.

[62] Koh, H. C. ; Chan, K. L. G. Data mining and customer relationship marketing in the banking industry, *Singapore Management Review*, 24, pages 1–28, 2002.

[63] Kuo, R. J.; Liao, J. L.; Tu, C. Integration of art neural network and genetic k means algorithm for analyzing web browsing paths in electronic commerce, *Decision Support Systems*, 40, pages 355–374, 2005.

[64] Lariviere, B.; Poel, D. V. D. Investigating the post-complaint period by means of survival analysis, *Expert Systems with Applications*, 29(3), pages 667–677, 2005.

[65] Lariviere, B.; Poel, D. Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling. The case of financial services, *Expert Systems with Applications*, 27(2), pages 277-285, 2004.

[66] Lee, M. H.; Choi, S. K.; Chung, G. H. Competition in Korea mobile telecommunications market: business strategy and regulatory environment, *Telecommunications Policy*, 25 (1), pages 124-138, 2001.

[67] Lu, J. X. Predicting customer churn in the telecommunications industry—An application of survival analysis modeling using SAS. In *Proc. SAS User Group International*, pages 27-33, 2002.

[68] Mallikarjuna, V., Mohan, G.; Kumar, D. Customer switching in mobile industry - an analysis of pre-paid mobile customers in AP circle of India. *International Journal of Research in Computer Application; Management*, 1(3), pages 63-66, 2011.

[69] Mattison, R. "*The Telco Churn Management Handbook*", XiT Press, Oakwood Hills, Illinois, 2005.

[70] Morik; Kopcke. Analyzing Customer Churn in Insurance Data, *Knowledge Discovery in Databases: PKDD 2004 Lecture Notes in Computer Science* 3202, pages 325-336, 2004

[71] Muata, K.; Bryson, O. Evaluation of Decision Trees: A Multi Criteria Approach, *Computers and Operational Research*, 31(11), pages 1933-1945, 2004.

[72] Mues, C.; Baesens, B.; Files, C.M.; Vanthienen, J. Decision Diagrams in Machine Learning: an Empirical Study on Real-life Credit-risk Data. *Expert Systems with Applications*, 27(2), pages 257–264 2004.

[73] Ng, K.; Liu, H. "Customer Retention via Data Mining." *Issues on the Application of data mining* pages 569-590, 2001.

[74] Ngai, E.W.T.; Xiu, Li; Chau, D.C.K. Application of data mining techniques in customer relationship management. A literature review and classification. *Expert Systems with Applications*, 36(2), pages 2592–2602, 2009.

[75] Nie, G.;Rowe, W.; Zhang, L.; Tian, Y. ; Shi, Y. Credit card churn forecasting by logistic regression and decision tree, *Expert Systems with Applications*,38, pages 15273–15285, 2011.

[76] Nord, J. H.; Nord, G. D. MIS research. Journal status and analysis, *Information and Management*, 29(1), pages 29–42, 1995.

[77] Omar, A.; Hossam, F.; Khalid, J.; Osama, H.; Nazeeh, G. Predicting Customer Churn in Telecom Industry using Multilayer Perceptron Neural Networks. *Modeling and Analysis*. Life Science Journal, 11(3), 2014.

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

[79] Phadke, C.; Uzunalioglu, H.; Mendiratta, V. B.; Kushnir, D.; Doran D. Prediction of Subscriber Churn Using Social Network Analysis, *Bell Labs Technical Journal*, 17(4), pages 63–75, 2013.

[80] Piotr, S. Global Perspectives Mobile Operator Customer Classification in Churn Analysis. *Technical University of Szczecin, Poland SAS Global Forum*, pages 1-5, 2008.

[81] Poku, K.; Zakari, M.; Sonali, A. Impact of Service Quality on Customer Loyalty in the Hotel Industry. An Empirical Study from Ghana. *International Review of Management and Business Research*, 22, pages 600-609, 2013.

[82] Qi, J.; Zhang, L.; Liu, Y.; Li, L.; Zhou, Y.; Shen, Y.; Liang, L. ; Li, H. AD Trees Logit model for customer churn prediction, *Ann Oper Res*, 168(1), pages 247–265, 2009.

[83] Rajkumar, P.; Rajkumar, H. Service Quality and Customers Preference of Cellular Mobile Service Providers. *Journal of Technology Management; Innovation*, 6(1), pages 38-45, 2011.

[84] Rashid, T. Classification of Churn and non-Churn Customers for Telecommunication Companies, *International Journal of Biometrics and Bioinformatics IJBB*, 3(5), 2008.

[85] Richter, Y.; Tov, Y.; Slonim, N. Predicting customer churn in mobile networks through analysis of social groups. *In Proc. SIAM International Conference on Data Mining*, 2010.

[86] Sarle, W. S. Neural Networks and Statistical Models, *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, 1994.

[87] Sathish, M.; Santhosh, K.;Naveen, K.J.; Jeevanantham, V. A. Study on Consumer Switching Behaviour in Cellular Service Provider. A Study with reference to Chennai. *Far East Journal of Psychology and Business*, 22, pages 71-81, 2011.

[88] Sayyed, M. A.; Tuteja, R.R. Data Mining Techniques, *International Journal of Computer Science and Mobile Computing IJCSMC,* 3(4), pages 879 – 883, 2014.

[89] Sharma, A.; Panigrahi, P.K. A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services. *International Journal of Computer Applications*, 27(11), pages 26-31, 2011.

[90] Singh, Y.; Chauhan, A.S. 2005. Neural Networks in Data Mining, *Journal of Theoretical and Applied Information Technology*, pages 37-42.

[91] Tamaddoni, M.; Moeini, I.; Akbari, A.; Akbarzadeh. A dual-step multi-algorithm approach for churn prediction in Pre-paid telecommunications service providers. The 6th International Conference on Innovation; Management, SÃO PAULO, Brazil, 2009.

[92] Torsten, D.; Martin, B. ; Krishnan, R. Estimating the effect of word of mouth on churn and cross-buying in the mobile phone market with Markov logic networks. *Decision Support Systems*, 51(3), pages 361-371, 2011.

[93] Umman, T.; Simsek, G. Customers churn analysis in telecommunication sector. *Istanbul University Journal of the School of Business Administration*, pages 35-49, 2010.

[94] Umman, T.; Serhat, G. Online shopping customer data using association rules and cluster analysis. *Advances in Data Mining. Applications and Theoretical Aspects,* Lecture Notes in Computer Science, 7987, pages127-136

[95] Verbeke, W.; Dejaeger, K.; Martens, D.; Hur, J.; Baesens, B. New insights into churn prediction in the telecommunication sector. A profit driven data mining approach, *Expert Systems with Applications*, 218(1), pages 211–229, 2012.

[96] Verbeke, W.; Martens, D.; Mues, C.; Baesens, B. Building comprehensible customer churn prediction models with advanced rule induction techniques. *Expert Systems with Applications*, 38(3) pages 2354-2364. doi:10.1016/j.eswa.2010.08.023, 2011.

[97] Vercellis, C. *Business Intelligence: Data Mining and Optimization for Decision Making*, John Wiley and Sons, Ltd. ISBN: 978-0-470-51138-1, 2009

[98] Wei, C. P.; Chiu, I. T. Turning telecommunications call details to churn prediction: A data mining approach. *Expert Systems with Applications*, 23(2), pages 103–112, 2002.

[99] Witten, I. H.; Frank, E. *Data Mining. Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, second edition, 2005.

[100] Wouter, V.; Karel, D.; David, M.; Joon, H.; Bart, B. New insights into churn prediction in the telecommunication sector. A profit driven data mining approach. *European Journal of Operational Research*, 218(1), pages 211–229, 2011.

[101] Xia, G. ; Jin, W. Model of Customer Churn Prediction on Support Vector Machine, *Systems Engineering Theory; Practice*, 28(1), 2008, pages 71-77, 2008

[102] Xie, Y.; Li, X.; Ngai, E. W. T. ; Ying, W. Customer churn prediction using improved balanced random forests, *Expert Systems with Applications*, 36 (3), pages 5445-5449, 2009.

[103] Zhu, C.; Qi, J., Wang, C. An experimental study on four models of customer churns prediction, Systems, Man and Cybernetics, 2009. SMC 2009. *IEEE International Conference*, pages 3199 –3204, 2009.