A SURVEY ON CHURN ANALYSIS

A PREPRINT

Jaehyun Ahn
Gyeonggi-do, Republic of Korea
jaehyunahn@sogang.ac.kr

October 30, 2020

ABSTRACT

In this paper, I present churn prediction techniques that have been released so far. Churn prediction is used in the fields of Internet services, games, insurance, and management. However, since it has been used intensively to increase the predictability of various industry/academic fields, there is a big difference in its definition and utilization. Our study can be used to select the definition of churn and its associated models suitable for the service field that researchers are most interested in by integrating fragmented churn studies in industry/academic fields.

Keywords
Churn Analysis · Churn Prediction · Machine Learning · Log Analysis · Data Mining · Customer Retention · Customer Relationship Management

1 Introduction

The term customer churn is commonly used to describe the propensity of customers who cease doing businesses with a company in a given time or contract [1]. Traditionally, studies on customer churn started from Customer Relation Management (CRM) [2]. It is crucial to prevent customer churn when operating services. In the past, the efficiency of customer acquisition relative to the number of churns was good. However, as the market saturated because of the globalization of services and fierce competition, customer acquisition costs rose rapidly [3] [4].

Reinartz, Werner, Jacquelyn S. Thomas, and Viswanathan Kumar. (2005) have shown that, for long-term business operations, putting efforts to increase the retention rate of all customers in terms of CRM is less efficient than putting efforts on a small number of targeted customer acquisition activities [22]. Similarly, Sasser, W. Earl. (1990) have suggested that retained customers generally return higher margins than randomly targeting new customers [23]. Additionally, Mozer, Michael C., et al. (2000) have proposed that, in terms of net return on investment, marketing campaigns for retaining existing customers are more efficient than putting efforts to attract new customers [16]. Reichheld et al. (1996) have shown that a 5 percent increase in customer retention rate achieved 35 percent and 95 percent increases in net present value of customers for a software company and an advertising agency, respectively [29]. As such, churn prediction can be used as a method to increase the retention rate of loyal customers and ultimately increase the value of the company.

Studies on customer churn have been proposed in various service fields. These studies on the churn analysis attempted to identify or predict in advance the likelihood that customers will churn using various indicators. The customer churn rate [5] is a typical customer churn analysis indicator. This refers to the ratio of subscribers who cancel a service to the total number of subscribers during a specific period [5] [32] [41]. The churn rate is the most widely used indicator for calculating the service retention period of subscribers in most service fields. Because of its importance and intuition, churn has been introduced in various service fields and developed to suit the characteristics of each field. Consequently, the research on the analysis of customer churn was fragmented according to each research field, thus the measurement criteria are all different. Currently, this is causing many problems. In the industry, communication costs arising from different churn criteria between service personnel in the process of fusing heterogeneous services (e.g., vehicle sharing service/insurance, online music service/department store) have been sharply increasing. Furthermore, since research on churn is simultaneously associated with two fields of engineering and business administration, it is not easy for researchers to describe two separate specialized fields on a single paper or to understand them.
In the past, customer churn of early days was used to define the customer’s status in the CRM. The CRM is a business management method that first emerged as a way of increasing the efficiency in areas of retail, marketing, sales, customer service, and supply-chain, and increasing efficiency and the customer value functions of the organization [2]. Since then, in the architectural point of view, the CRM has evolved and become divided into operational CRM and analytical CRM. The analytical CRM is focused on developing databases and resources containing customer characteristics and attitudes [24] [25] [30]. The analytical CRM has been initially used for creating appropriate marketing strategies using customer status and customer behavior data, and particularly, it has been used to fulfill the individual and unique needs of customers [26]. From this point on, IT and knowledge management related technologies have been utilized, and companies started applying dedicated technologies for acquisition, retention, churn, and selection of customers [27], and ever since the technologies of IT field became implemented in the CRM, various companies began to use such technologies in business areas including data warehouse, website, telecommunication, and banking [28]. As described earlier, with studies on CRM claiming that increasing the retention rate of small number of existing customers is more efficient than acquiring new customers, churn analysis has become one of the important personalized customer management techniques [20] [28] [59]. There were survey papers that collected and summarized churn analysis techniques in the telecommunications field [6] [7] [8] [9]. However, these studies are limited to the telecommunications field, and the log data used for the churn analysis do not include time series features, retention and survival, and KPI (Key Performance Indicator) features. There were also papers applied to services using various deep learning model-based churn analysis techniques in terms of computer science [10] [11]. However, these studies are limited to the deep learning algorithm, and lack underlying models and parameter description. There are also a few survey papers on churn, yet they do not cover the latest deep learning techniques but cover only churn in specific industrial fields [12] [13]. The trend of building churn prediction models is changing, and performance is rapidly improving. However, because of fragmented previous studies, there are many difficulties for researchers to launch new research on churn. In order to address these issues, this survey paper describes the differences in the definition of churn prediction algorithms in the fields of business administration, marketing, IT, telecommunications, newspaper publishing, insurance, and psychology, and compares differences in churn loss and feature engineering. In addition, I classify and explain the cases of churn prediction models based on this. Our study provides classification information for more detailed technologies on churn in a wider range than previous survey papers. Our research can reduce confusion about the churn criteria that are being fragmented and utilized across multiple industry/academic fields, and can be of a practical help in applying them to prediction models. In particular, this paper presents a deep learning model among machine learning techniques designed to solve non-contractual customer churns, which have recently appeared with the advancement of industries. The structure of this paper is as follows. Chapter 2 introduces typical definitions for churn in each business domain and their differences. Chapter 3 presents churn application cases in various business domains.

2 Definition of Churn

Churn has been defined in various ways in multiple industries. In this chapter, I describe two typical types. As seen in Table 1, typical papers with different criteria for defining churn are summarized. In general, the dictionary definition of churn is known as the prolonged period of inactivity [14]. However, the criteria for ‘inactivity’ and ‘prolonged’ are different according to each research field. Such inconsistency is frequently found due to more services of modern days adopting loose subscription terms because of competition. In the past, customer churn had occurred explicitly through contractual cancellations, however, in the modern services including Internet and retail services, frequent customer churns occur due to the low customers’ investment costs [19] [20] [21] [96]. These non-contractual customer churns occur due to low switching cost for changing the service [31]. Thus, I was divide the criteria of churn into contractual churn and non-contractual churn. Descriptions of each churn is as follows.

The first criterion is contractual churn. Contractual churn refers to churn that a customer does not extend the contract even when the contract renewal date is reached [15] [16]. This churn means that a customer loses interest in the relevant service area and changes his/her position to a state where re-entry is no longer possible. It is usually present in churn problems occurring when customers close their banking accounts or when switching their carrier operator from one service to another. In addition, contractual churn is frequently found in a flat-rate service such as music and movie streaming services.

The second criterion is non-contractual churn. In general, in a non-contractual situation, customers can leave the service/contract without time constraints. In the service operating perspective, a criterion for churn is first constructed, then a customer that meets such criterion is categorized as the churn customer. To conduct this, the customer’s behavioral changed date is counted [96]. When this inactivity or behavioral changed period exceeds the threshold, the customer is regarded as a churn customer. During this process, the period that is set as the threshold of the inactivity date is called the time window [17]. The defining of non-contractual churn has made it possible to infer the probability of the customers who are likely to churn within the certain period. The time window method is frequently used when
Table 1: Customer churn decision by period and method

| Setting          | Churn observation criteria | Publishing Information |
|------------------|----------------------------|------------------------|
| Contractual      | Monthly                    | Chen, Yan, et al. (2018) [15] |
| Contractual      | Monthly                    | Mozer, Michael C., et al. (2000) [16] |
| Contractual      | Monthly                    | Dahiya, Kiran, and Surbhi Bhatta. (2015) [65] |
| Contractual      | Monthly                    | Bahnisen, Alejandro Correa, Djamila Aouada, and Björn Ottersten. (2015) [66] |
| Contractual      | Monthly                    | Coussenheim, Kristof, and Dirk Van den Poel. (2008) [67] |
| Contractual      | Monthly                    | Radosavljevik, Dejan, Peter van der Putten, and Kim Kyllesbech Larsen. (2010) [68] |
| Contractual      | Monthly                    | Gladý, Nicolas, Bart Baesens, and Christophe Croux. (2009) [38] |
| Contractual      | Monthly                    | Hung et al. (2006) [32] |
| Contractual      | Monthly                    | Burez and van den Poel (2007) [131] |
| Contractual      | Monthly                    | Burez and van den Poel (2008) [132] |
| Contractual      | Monthly                    | Burez, Jonathan, and Dirk Van den Poel. (2009) [95] |
| Contractual      | Monthly                    | Madden, Savage, and Coble-Neal (1999) [133] |
| Contractual      | Monthly                    | Gerpott et al. (2001) [134] |
| Contractual      | Monthly                    | Seo et al. (2008) [135] |
| Contractual      | Monthly                    | Pendharkar (2009) [136] |
| Contractual      | Monthly                    | Chu, Tsai, and Ho (2007) [137] |
| Contractual      | Monthly                    | Kim and Yoon (2004) [139] |
| Contractual      | Monthly                    | Aih et al. (2006) [33] |
| Contractual      | Monthly                    | Wet, Chih-Ping, and I-Tang Chiu. (2002) [41] |
| Contractual      | Monthly                    | Neslin, Scott A., et al. (2006) [54] |
| Contractual      | Binary                     | Zhang, Rong, et al. (2017) [11] |
| Contractual      | Binary                     | Dechant, Andrea, Martin Spann, and Jan U. Becker. (2019) [46] |
| Contractual      | Binary                     | Xie et al. (2009) [91] |
| Contractual      | Binary                     | Athanasopoulos (2000) [138] |
| Non-contractual  | Monthly                    | Larivière, Bart, and Dirk Van den Poel. (2004) [36] |
| Non-contractual  | Daily                      | Lee, Eunjo, et al. (2018) [10] |
| Non-contractual  | Daily                      | Lee, Eunjo, et al. (2018) [17] |
| Non-contractual  | Daily                      | Tamaddoni Jahromi, Ali, et al. (2010) [20] |
| Non-contractual  | Daily                      | Hadiji, Fabian, et al. (2014) [42] |
| Non-contractual  | Daily                      | Yang, Wanshan, et al. (2019) [43] |
| Non-contractual  | Binary                     | Buckinx, Wouter, and Dirk Van den Poel. (2005) [96] |

Customer churn can be defined differently for various service characteristics. Generally, customer churn can be decided by customer’s contract revocation, non-use time window period, and service deletion.

Analyzing activity logs these days in a non-contractual situation. When a customer does not use a service for a certain period of time, this method regarding customer as churned. Internet services do not usually delete accounts. Therefore, the Internet service interprets the log-in as prolonged, that is, the retention of the service, and interprets unconnected access for a certain period of time as churn [17]. Fig. 1 schematically illustrates the non-contractual churn case with time window method. The log of Fig. 1 was recorded for 10 weeks. The time window is set to 4 weeks from Week 4 to Week 7. Six users in Fig. 1 showed their activities in each week, and their activities were logged. In the time window period, users A, B, and C without any activity logs from Week 4 to Week 7 are regarded as churn, and the other users D, E, and F with activity logs are regarded as retention.

Churn analysis is usually performed to improve business outcomes. Therefore, in most churn prediction problems, the churn period is defined as a section that can restore customers’ trust. If the time period during which a customer completely churns is selected as a time window, the period for churn definition exponentially increases and it does not provide any gain in terms of business as changing the will of the customers who want to churn is deemed impossible [17]. The contractual mentioned above are close to customers’ complete churn from a service. Therefore, these days majority of the log-based churn prediction problems use the probabilistic method to determine whether customers are churning or not and to give customers incentive to reuse their service.
The criteria for setting the time window are different for each service feature. Yang, Wanshan, et al. (2019) analyzed log data to define the churn period of mobile games, and the analysis result showed that more than 95% of customers did not return when they were absent for 3 consecutive days. They set 3 days as the time window churn period [43]. Lee, Eunjo, et al. (2018) defined the period during which 75% of customers were continuously unconnected as a churn section by taking into consideration the characteristics of PC game services [17]. After collecting customers’ unconnected periods, they drew a cumulative data graph. They selected the section where more than 75% of customers churned as the time window. Fig. 2 schematically illustrates the cumulative data of consecutive unconnected days collected by Lee, Eunjo, et al. (2018). According to Fig. 2, the period during which customers were unconnected more than 75% was 14 weeks. Therefore, the time window churn period is 14 weeks.

As described above, there are two customer churn types, which are contractual churn and non-contractual churn. Additionally, there are three churn observation criteria as follows: monthly, daily, and binary. The monthly and daily churn observations are related to the cycle in which the customer’s status is updated in the database. The binary churn
observation is acquired by manipulating this database. In general, binary churn is determined by the existence of contract in the contractual settings. In the non-contractual settings, the company defines the customer inactivity features, and when a customer meets the inactivity or disloyal customer feature, the customer is regarded as binary churn [96]. The reason for having multiple ways of defining customer churn is to periodically monitor the customers’ status changes. And through such observation, the expected net business value can be increased by predicting customer churn rates and providing possible churn customers with incentives to retain them from leaving [16] [23].

3 Churn Analysis in Various Business Field

The majority of the early studies on churn were conducted from a management perspective, especially CRM (Customer Relation Management) [30] [31]. CRM churn covers all churn problems that can occur in the process of customer identification, customer attraction, customer retention, and customer development. Modern churn prediction problems are mainly analyzed using log data. A log is trace data that remains when using Internet services. Therefore, the churn prediction models implemented using log data can be used for Internet services in various industries. There are 12 business fields that performed churn prediction.

The telecommunications industry accounts for the majority of previous studies on churn. Telecommunications services have high customer stickiness despite high customer acquisition costs. Therefore, if customer churn is prevented and appropriate incentives are provided, it is of great help in maintaining sales [16] [32] [33] [34].

The financial and insurance industries also predict customer churn. Zhang, Rong, et al. (2017) stressed the need to build churn prediction models and prevent churn, referring to high customer acquisition costs and high customer values in the insurance industry [11]. Chiang, Ding-An, et al. (2003) mentioned that customer values were high in the online financial market, and created a churn scenario according to the financial product selection and customers’ financial product selection sequence using the Apriori algorithm [35]. Larivières, Bart, and Dirk Van den Poel. (2004), based on the assumption that the customer group was different according to the financial product attribute, demonstrated that the likelihood of churn differed depending on the tendency of customers who selected financial products by measuring the survival time for each product [36]. Zopounidis, Constantin, Maria Mavri, and George Ioannou. (2008) measured the switching rate of financial products, and the survival period of customers for each product to discover attractive products [37]. Here, as the survival period is short, churn occurs more frequently, which is used as an indicator to measure the need to supplement financial products. Glady, Nicolas, Bart Baesens, and Christophe Croux. (2009) measured the customer lifetime values and the decrease in expected earnings over time as an indicator corresponding to customer loyalty [38]. During this process, machine learning was used to calculate the churn rate which was used to estimate the customer lifetime values.

Later on, studies on churn have been actively conducted in the gaming field as in the telecommunication field. These services have a fast cycle of customer inflow and churn because of mass competition. However, if a single service is run for a long time, the service competition intensifies and the Customer Acquisition Cost (CAC) tends to increase [16] [39] [128]. As the CAC gets larger, the technology to predict and prevent churn becomes more crucial. Viljanen, Markus, et al. (2016) applied the survival analysis to mobile games and calculated the churn rate, similar to the churn prediction of financial services [40]. The game sector actively uses machine learning techniques when conducting research on churn because of the large volume of log data [10] [42] [43]. Milošević, Miloš, Nenad Živić, and Igor Andjelković. (2017) created a model predicting churn in the study on game churn, gave churn prevention incentives by finding out and dividing probable churn customers into A/B groups, and demonstrated actual effects statistically [44]. Runge, Julian, et al. (2014) conducted a similar study, and revealed that existing customers with a high possibility of churn had a higher marketing response rate when compared to general marketing targets [45].

Furthermore, the music streaming service field even held a competition to build a prediction model, and research on churn was also conducted in the Internet service and newspaper subscription fields. The newspaper subscription and music streaming service offer fixed-rate services, and customer churn is consistent with the contract renewal period. On the other hand, because the Internet service goes into an inactive state as customers wish, contract renewal takes place nearly-real-time. Research on churn prediction was also conducted in online dating, online commerce, Q&A services, and social network-based services [46].

There were some studies which approached customer churn from a psychological perspective. Borbora, Zoheb, et al. (2011) analyzed that customers churned when their motivation to use games changed by combining the motivation theory with customers using MMO RPG games [47]. Yee, Nick. (2016) surveyed approximately 250,000 gamers, and showed that customers’ attitudes toward games were clustered by country, race, and age [48].
In the marketing field, Glady, Nicolas, Bart Baesens, and Christophe Croux. (2009) used the features from a marketing perspective such as RFM (Recency, Frequency and Monetary) and CLV (Customer Life time Value) for churn prediction [38].

Studies on churn prediction were conducted in the human resources and energy fields although they were minority. Saradhi, V. Vijaya, and Girish Keshav Palshikar. (2011) conducted research on churn to reduce retraining costs when employees churned and to prove employee value in the human resources field [50]. Moeyersoms, Julie, and David Martens. (2015) estimated whether customers would churn to another energy supplier based on energy data and socio-demographic data provided to customers [51].

4 Conclusions

In this study, I compared the churn prediction analysis techniques using log data. Churn analysis is used in the fields of Internet services and games, insurance, and management. Research on churn prediction usually begins to improve business outcomes. Therefore, the time window is used to select potential churning customers rather than measuring a customer’s complete churn. Loss costs for customer churn are calculated by CAC or CLV. In the past, when predicting customer churn, researchers used survival analysis or time series analysis using statistics, graph theory, and traditional machine learning algorithms. Churn prediction analysis using deep learning algorithms has recently emerged. Deep learning algorithms have been found to outperform other algorithms. This is likely due to large quantities of customer log data being collected via computers and the churn prediction model utilizing the entire set of this acquired data to make churn predictions. Churn Prediction Models of this paper used deep learning for churn prediction with data timestamps in the order of seconds or with vast amounts of customer log data in total. In such case, feature engineering techniques for processing logs have a significant effect on model performance enhancement. Unlike other modeling techniques, the deep learning model is capable of converting high dimensional sparse log data into low dimensional dense features by embedding time series features. Also the deep learning model can learn customer’s behavioral patterns from vast amount of data by layer-wise stacked neurons structure. Therefore, given minute timestamps and abundant observations, applying this data to deep learning algorithms for the generation of latent features is expected to produce better performance than conventional churn prediction models.

This is because as the log data in these days is collected for a longer period and deep learning algorithm get an advantage to catch customers’ latent status compared with older algorithms. In other words, the reason deep learning algorithms are receiving spotlight today is due to the vast amount of data used in modern churn predictions, and its ability to capture minute changes. As mentioned earlier in the text, traditional churn prediction algorithms including statistics methods are still actively used today. This is due to variations in which churn prediction model has the best performance depending on the data format. Churn prediction models using deep learning is a new solution with a good structure for predicting modern churn datasets. Therefore, to solve the problem at hand, readers will need to understand the format of the churn dataset and apply a suitable algorithm to solve the churn prediction problem.

Furthermore, I also outlined a performance evaluation method for comparing the various churn prediction algorithms used from the past to the present. Most churn prediction models are related to customer relation management. For example, there may be performance differences depending on whether the churn prediction model is robust against false positives or false negatives. According to the research of this paper, many articles use AUC as a performance measurement method aside from standard precision. In general, as there are fewer churn customers than non-churn customers, a performance specific method focused on churn customers will be needed. The ROC curve is a graph of the rate at which the model correctly predicts churn customers and the rate at which residual customers are predicted to be churn customers. Therefore, it is a performance measurement method that focuses on the prediction of churn customers. In this study, I comprehensively compared the churn prediction problems. This paper helps to find a method that meets the needs of researchers among various churn prediction algorithms. Furthermore, this paper is expected to be used to improve services and build better churn analysis models.

References

[1] Chandar, M., Arijit Laha, and P. Krishna. “Modeling churn behavior of bank customers using predictive data mining techniques.” National conference on soft computing techniques for engineering applications (SCT-2006). 2006.

[2] Parvatiyar, Atul, and Jagdish N. Sheth. “Customer relationship management: Emerging practice, process, and discipline.” Journal of Economic and Social Research 3.2 (2001).

[3] Fields, Tim. “Mobile & social game design: Monetization methods and mechanics”, CRC Press, 2014.
[4] Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. “Profit optimizing customer churn prediction with Bayesian network classifiers.” *Intelligent Data Analysis* 18.1 (2014): 3-24.

[5] Investopedia: https://www.investopedia.com/terms/c/churnrate.asp (Access on 4 May 2020)

[6] Almana, Amal M., Mehmet Sabih Aksoy, and Rasheed Alzahrani. “A survey on data mining techniques in customer churn analysis for telecom industry.” *International Journal of Engineering Research and Applications* 45 (2014): 165-171.

[7] Ahmed, Ammara, and D. Maheswari Linen. “A review and analysis of churn prediction methods for customer retention in telecom industries.” 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2017.

[8] Vafeiadis, Thanasis, et al. “A comparison of machine learning techniques for customer churn prediction.” *Simulation Modelling Practice and Theory* 55 (2015): 1-9.

[9] Ahmed, Mehreen, et al. “A survey of evolution in predictive models and impacting factors in customer churn.” *Advances in Data Science and Adaptive Analysis* 9.03 (2017): 1750007.

[10] Lee, Eunjo, et al. “Game data mining competition on churn prediction and survival analysis using commercial game log data.” *IEEE Transactions on Games* 11.3 (2018): 215-226.

[11] Zhang, Rong, et al. “Deep and shallow model for insurance churn prediction service.” 2017 IEEE International Conference on Services Computing (SCC). IEEE, 2017.

[12] García, David L., Ángela Nebot, and Alfredo Vellido. “Intelligent data analysis approaches to churn as a business problem: a survey.” *Knowledge and Information Systems* 51.3 (2017): 719-774.

[13] Mohammed, et al. “Customer Churn in Mobile Markets: A Comparison of Techniques.” *International Business Research* 8.6 (2015).

[14] Periáñez, África, et al. “Churn prediction in mobile social games: Towards a complete assessment using survival ensembles.” 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2016.

[15] Chen, Yian, et al. “Wsdm cup 2018: Music recommendation and churn prediction.” *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 2018.

[16] Mozer, Michael C., et al. “Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry.” *IEEE Transactions on neural networks* 11.3 (2000): 690-696.

[17] Lee, Eunjo, et al. “Profit Optimizing Churn Prediction for Long-term Loyal Customer in Online games.” *IEEE Transactions on Games* (2018).

[18] Evermann, Joerg, Jana-Rebecca Rehse, and Peter Fettke. “Predicting process behavior using deep learning.” *Decision Support Systems* 100 (2017): 129-140.

[19] Ma, Shaohui. “On Optimal Time for Customer Retention in Non-Contractual Setting.” Available at SSRN 1529284 (2009).

[20] Tamaddoni Jahromi, Ali, et al. “Modeling customer churn in a non-contractual setting: the case of telecommunications service providers.” *Journal of Strategic Marketing* 18.7 (2010): 587-598.

[21] Nielsen, A.C. “Major study to track store switching” *Retail World* (2009)

[22] Reinartz, Werner, Jacquelyn S. Thomas, and Viswanathan Kumar. “Balancing acquisition and retention resources to maximize customer profitability.” *Journal of marketing* 69.1 (2005): 63-79.

[23] Sasser, W. Earl. “Zero defections: quality comes to services.” *Harvard Business Review* 68.5 (1990): 105-111.

[24] He, Zengyou, et al. “Mining class outliers: concepts, algorithms and applications in CRM.” *Expert Systems with applications* 27.4 (2004): 681-697.

[25] Teo, Thompson SH, Paul Devadoss, and Shan L. Pan. “Towards a holistic perspective of customer relationship management (CRM) implementation: A case study of the Housing and Development Board, Singapore.” *Decision support systems* 42.3 (2006): 1613-1627.

[26] Shaw, Michael J., et al. “Knowledge management and data mining for marketing.” *Decision support systems* 31.1 (2001): 127-137.

[27] Komenar, Margo. “Electronic marketing”. John Wiley & Sons, Inc., 1996.

[28] Bose, Ranjit. “Customer relationship management: key components for IT success.” *Industrial management & Data systems* (2002).

[29] Reichheld, Frederick F., Thomas Teal, and Douglas K. Smith. “The loyalty effect.” (1996): 78-84.
[30] Ngai, Eric WT, Li Xiu, and Dorothy CK Chau. “Application of data mining techniques in customer relationship management: A literature review and classification.” Expert systems with applications 36.2 (2009): 2592-2602.

[31] Lejeune, Miguel APM. “Measuring the impact of data mining on churn management.” Internet Research: Electronic Networking Applications and policy, 11, 375-387 (2001).

[32] Hung, Shin-Yuan, David C. Yen, and Hsiu-Yu Wang. “Applying data mining to telecom churn management.” Expert Systems with Applications 31.3 (2006): 515-524.

[33] Ahn, Jae-Hyeon, Sang-Pil Han, and Yung-Seop Lee. “Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry.” Telecommunications policy 30.10-11 (2006): 552-568.

[34] Au, Wai-Ho, Keith CC Chan, and Xin Yao. “A novel evolutionary data mining algorithm with applications to churn prediction.” IEEE transactions on evolutionary computation 7.6 (2003): 532-545.

[35] Chiang, Ding-An, et al. “Goal-oriented sequential pattern for network banking churn analysis.” Expert Systems with Applications 25.3 (2003): 293-302.

[36] Larivièere, Bart, and Dirk Van den Poel. “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 (2004): 277-285.

[37] Zopounidis, Constantin, Maria Mavri, and George Ioannou. “Customer switching behavior in Greek banking services using survival analysis.” Managerial Finance (2008).

[38] Gladys, Nicolas, Bart Baesens, and Christophe Croux. “Modeling churn using customer lifetime value.” European Journal of Operational Research 197.1 (2009): 402-411.

[39] Xia, Guo-en, and Wei-dong Jin. “Model of customer churn prediction on support vector machine.” Systems Engineering-Theory & Practice 28.1 (2008): 71-77

[40] Viljanen, Markus, et al. “Modelling user retention in mobile games.” 2016 IEEE Conference on Computational Intelligence and Games (CIG). IEEE, 2016.

[41] Wei, Chih-Ping, and I-Tang Chiu. “Turning telecommunications call details to churn prediction: a data mining approach.” Expert systems with applications 23.2 (2002): 103-112.

[42] Hadjii, Fabian, et al. “Predicting player churn in the wild.” 2014 IEEE Conference on Computational Intelligence and Games. IEEE, 2014.

[43] Yang, Wanshan, et al. “Mining Player In-game Time Spending Regularity for Churn Prediction in Free Online Games.” 2019 IEEE Conference on Games (CoG). IEEE, 2019.

[44] Milošević, Miloš, Nenad Živić, and Igor Andjelković. “Early churn prediction with personalized targeting in mobile social games.” Expert Systems with Applications 83 (2017): 326-332.

[45] Runge, Julian, et al. “Churn prediction for high-value players in casual social games.” 2014 IEEE conference on Computational Intelligence and Games. IEEE, 2014.

[46] Dechant, Andrea, Martin Spann, and Jan U. Becker. “Positive customer churn: An application to online dating.” Journal of Service Research 22.1 (2019): 90-100.

[47] Borbora, Zoheb, et al. “Churn prediction in mmorpgs using player motivation theories and an ensemble approach.” 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. IEEE, 2011.

[48] Yee, Nick. “The gamer motivation profile: What we learned from 250,000 gamers.” 2016 Annual Symposium on Computer-Human Interaction in Play, 2016.

[49] Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee. “RFM and CLV: Using iso-value curves for customer base analysis.” Journal of marketing research 42.4 (2005): 415-430.

[50] Saradh, V. Vijaya, and Girish Keshav Palshikar. “Employee churn prediction.” Expert Systems with Applications 38.3 (2011): 1999-2006.

[51] Moeyersoms, Julie, and David Martens. “Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector.” Decision support systems 72 (2015): 72-81.

[52] Sifa, Rafet, et al. “Predicting purchase decisions in mobile free-to-play games.” Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference, 2015.

[53] Richter, Yossi, Elad Yom-Tov, and Noam Slonim. “Predicting customer churn in mobile networks through analysis of social groups.” Proceedings of the 2010 SIAM international conference on data mining. Society for Industrial and Applied Mathematics, 2010.
[54] Neslin, Scott A., et al. “Defection detection: Measuring and understanding the predictive accuracy of customer churn models.” *Journal of marketing research* 43.2 (2006): 204-211.

[55] Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. “A novel profit maximizing metric for measuring classification performance of customer churn prediction models.” *IEEE transactions on knowledge and data engineering* 25.5 (2012): 961-973.

[56] Sifa, Rafet, Christian Bauckhage, and Anders Drachen. “The Playtime Principle: Large-scale cross-games interest modeling.” *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 2014.

[57] Dror, Gideon, et al. “Churn prediction in new users of Yahoo! answers.” *Proceedings of the 21st International Conference on World Wide Web*. 2012.

[58] Nimmagadda, Sravya, Akshay Subramaniam, and Man Long Wong. “Churn prediction of subscription user for a music streaming service.” (2017).

[59] Ngai, Eric WT. “Customer relationship management research (1992 -2002).” *Marketing intelligence & planning* (2005).

[60] Breiman, Leo. “Statistical modeling: The two cultures (with comments and a rejoinder by the author).” *Statistical science* 16.3 (2001): 199-231.

[61] Matthew Stewart. “The Actual Difference Between Statistics and Machine Learning” *Towards Data Science*, https://towardsdatascience.com/the-actual-difference-between-statistics-and-machine-learning-64b49f07ea3, May 25th (2019).

[62] Bzdok, D., Altman, N. and Krzywinski, M. “Statistics versus machine learning.” *Nat Methods* 15, 233–234 (2018). https://doi.org/10.1038/nmeth.4642

[63] Witten, Ian H., and Eibe Frank. “Data mining: practical machine learning tools and techniques with Java implementations.” *Acm Sigmod Record* 31.1 (2002): 76-77.

[64] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[65] Dahiy, Kiran, and Surbhi Bhatia. “Customer churn analysis in telecom industry,” *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*. IEEE, 2015.

[66] Bahnsen, Alejandro Correa, Djamila Aouada, and Björn Ottersten. “A novel cost-sensitive framework for customer churn predictive modeling.” *Decision Analytics* 2.1 (2015): 5.

[67] Coussement, Kristof, and Dirk Van den Poel. “Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques.” *Expert systems with applications* 34.1 (2008): 313-327.

[68] Radosavljevik, Dejan, Peter van der Putten, and Kim Kyllesbech Larsen. “The impact of experimental setup in prepaid churn prediction for mobile telecommunications: What to predict, for whom and does the customer experience matter?” *Trans. MLDM* 3.2 (2010): 80-99.

[69] Kim, Seungwook, et al. “Churn prediction of mobile and online casual games using play log data.” *PloS one* 12.7 (2017).

[70] Bindewald, Jason M., Gilbert L. Peterson, and Michael E. Miller. “Clustering-based online player modeling.” *Computer Games*. Springer, Cham, 2016. 86-100.

[71] Ben Lewis-Evans. “Finding Out What They Think: A Rough Primer To User Research, Part 2” *Gamasutra*, May 15th (2012).

[72] Drachen, Anders, Magy Seif El-Nasr, and Alessandro Canossa, eds. “Game Analytics: Maximizing the Value of Player Data.” *Springer*, 2013.

[73] Bertens, Paul, Anna Guitart, and África Periáñez. “Games and big data: A scalable multi-dimensional churn prediction model.” *2017 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 2017.

[74] Nozhnin, Dmitry. “Predicting churn: When do veterans quit.” *Gamasutra*. August 30th (2012).

[75] Tamassia, Marco, et al. “Predicting player churn in destiny: A Hidden Markov models approach to predicting player departure in a major online game.” *2016 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 2016.

[76] Bastiaan van der Palen. “Predicting player churn using game-design-independent features across casual free-to-play games” *Tilburg School of Humanities*. http://arno.uvt.nl/show.cgi?fid=144997, 2017.

[77] Fields, Tim, and Brandon Cotton. “Social game design: Monetization methods and mechanics.” *CRC Press*, 2011.
[78] Kawale, Jaya, Aditya Pal, and Jaideep Srivastava. “Churn prediction in MMORPGs: A social influence based approach.” 2009 International Conference on Computational Science and Engineering. Vol. 4. IEEE, 2009.

[79] Kristensen, Jeppe Theiss, and Paolo Burelli. “Combining Sequential and Aggregated Data for Churn Prediction in Casual Freemium Games.” 2019 IEEE Conference on Games (CoG). IEEE, 2019.

[80] Guitart, Anna, Pei Pei Chen, and África Periáñez. “The Winning Solution to the IEEE CIG 2017 Game Data Mining Competition.” Machine Learning and Knowledge Extraction 1.1 (2019): 252-264.

[81] Viljanen, Markus, et al. “A/B-test of retention and monetization using the Cox model.” Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference. 2017.

[82] Lee, eunjo. “User behavior modeling in online games using machine learning techniques” June 2019. Korea University Graduate School of Information Security. [UCI] I804:11009-00000084744. (2019)

[83] Bauckhage, Christian, et al. “How players lose interest in playing a game: An empirical study based on distributions of total playing times.” 2012 IEEE Conference on Computational Intelligence and Games (CIG). IEEE, 2012.

[84] Shores, Kenneth B., et al. “The identification of deviance and its impact on retention in a multiplayer game.” Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. 2014.

[85] Debeauvais, Thomas, et al. “If you build it they might stay: retention mechanisms in World of Warcraft.” Proceedings of the 6th International Conference on Foundations of Digital Games. 2011.

[86] Castro, Emiliano G., and Marcos SG Tsuzuki. “Churn prediction in online games using players’ login records: A frequency analysis approach.” IEEE Transactions on Computational Intelligence and AI in Games 7.3 (2015): 255-265.

[87] Wolters, Hans, Jim Baer, and Girish Keswani. “Method to detect and score churn in online social games.” U.S. Patent No. 8,790,168. 29 Jul. 2014.

[88] Coussement, Kristof, and Koen W. De Bock. “Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning.” Journal of Business Research 66.9 (2013): 1629-1636.

[89] Butgereit, Laurie. “Work Towards Using Micro-services to Build a Data Pipeline for Machine Learning Applications: A Case Study in Predicting Customer Churn.” 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE). IEEE, 2020.

[90] Wright, Christine. “System and method for predicting and preventing customer churn.” U.S. Patent Application No. 10/419,463.

[91] Xie, Yaya, et al. “Customer churn prediction using improved balanced random forests.” Expert Systems with Applications 36.3 (2009): 5445-5449.

[92] Anil Kumar, Dudyala, and Vadlamani Ravi. “Predicting credit card customer churn in banks using data mining.” International Journal of Data Analysis Techniques and Strategies 1.1 (2008): 4-28.

[93] Nie, Guangli, et al. “Credit card churn forecasting by logistic regression and decision tree.” Expert Systems with Applications 38.12 (2011): 15273-15285.

[94] Ali, Özden Gür, and Umut Arttürk. “Dynamic churn prediction framework with more effective use of rare event data: The case of private banking.” Expert Systems with Applications 41.17 (2014): 7889-7903.

[95] Burez, Jonathan, and Dirk Van den Poel. “Handling class imbalance in customer churn prediction.” Expert Systems with Applications 36.3 (2009): 4626-4636.

[96] Buckinx, Wouter, and Dirk Van den Poel. “Customer base analysis: partial defection of behaviorally loyal clients in a non-contractual FMCG retail setting.” European journal of operational research 164.1 (2005): 252-268.

[97] Clemente, M., V. Giner-Bosch, and S. San Matías. “Assessing classification methods for churn prediction by composite indicators.” Manuscript, Dept. of Applied Statistics, OR & Quality, Universitat Politècnica de València, Camino de Vera s/n 46022 (2010).

[98] Morik, Katharina, and Hanna Köpcke. “Analysing customer churn in insurance data—a case study.” European conference on principles of data mining and knowledge discovery. Springer, Berlin, Heidelberg, 2004.

[99] Hur, Yeon, and Sehun Lim. “Customer churning prediction using support vector machines in online auto insurance service.” International Symposium on Neural Networks. Springer, Berlin, Heidelberg, 2005.

[100] Risselada, Hans, Peter C. Verhoef, and Tammo HA Bijmolt. “Staying power of churn prediction models.” Journal of Interactive Marketing 24.3 (2010): 198-208.

[101] Coussement, Kristof, Dries F. Benoit, and Dirk Van den Poel. “Improved marketing decision making in a customer churn prediction context using generalized additive models.” Expert Systems with Applications 37.3 (2010): 2132-2143.
[102] Tamaddoni, Ali, Stanislav Stakhovych, and Michael Ewing. “Comparing churn prediction techniques and assessing their performance: a contingent perspective.” Journal of service research 19.2 (2016): 123-141.

[103] De Bock, Koen W., and Dirk Van den Poel. “An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction.” Expert Systems with Applications 38.10 (2011): 12293-12301.

[104] Chen, Min. “Music Streaming Service Prediction with MapReduce-based Artificial Neural Network.” 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2019.

[105] Ngonmang, Blaise, Emmanuel Viennet, and Maurice Tchuente. “Churn prediction in a real online social network using local community analysis.” 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2012.

[106] Yu, Xiaobing, et al. “An extended support vector machine forecasting framework for customer churn in e-commerce.” Expert Systems with Applications 38.3 (2011): 1425-1430.

[107] Dalvi, Preeti K., et al. “Analysis of customer churn prediction in telecom industry using decision trees and logistic regression.” 2016 Symposium on Colossal Data Analysis and Networking (CDAN). IEEE, 2016.

[108] Ge, Yizhe, et al. “Customer Churn Analysis for a Software-as-a-service Company.” 2017 Systems and Information Engineering Design Symposium (SIEDS). IEEE, 2017.

[109] Baumann, Annika, et al. “Maximize What Matters: Predicting Customer Churn With Decision-Centric Ensemble Selection.” 2015 European Conference on Information Systems (ECIS). 2015.

[110] Dasgupta, Koustuv, and Kwok-Leung Tsui. “Churn detection via customer profile modelling.” International Journal of Production Research 44.14 (2006): 2913-2933.

[111] Hadden, John, et al. “Churn prediction: Does technology matter.” International Journal of Intelligent Technology 1.2 (2006): 104-110.

[112] Nath, Shyam V., and Ravi S. Behara. “Customer churn analysis in the wireless industry: A data mining approach.” Proceedings-annual meeting of the decision sciences institute. Vol. 561. 2003.

[113] Lemmens, Aurélie, and Christophe Croux. “Bagging and boosting classification trees to predict churn.” Decision Support Systems 51.3 (2011): 361-371.

[114] Idrees, Adnan, Muhammad Rizwan, and Asifullah Khan. “Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies.” Computers & Electrical Engineering 38.6 (2012): 1808-1819.

[115] Idris, Adnan, Muhammad Rizwan, and Asifullah Khan. “Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies.” Computers & Electrical Engineering 38.6 (2012): 1808-1819.

[116] Huang, Ying, and Tahar Kechadi. “An effective hybrid learning system for telecommunication churn prediction.” Expert Systems with Applications 40.14 (2013): 5635-5647.

[117] Huang, Bingquan, Mohand Tahar Kechadi, and Brian Buckley. “Customer churn prediction in telecommunications.” Expert Systems with Applications 39.1 (2012): 1414-1425.

[118] Dierkes, Torsten, Martin Bichler, and Ramayya Krishnan. “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 (2011): 361-371.

[119] Kim, Kyoungok, Chi-Hyuk Jun, and Jaewook Lee. “Improved churn prediction in telecommunication industry by analyzing a large network.” Expert Systems with Applications 41.15 (2014): 6575-6584.
[125] Keramati, Abbas, et al. “Improved churn prediction in telecommunication industry using data mining techniques.” *Applied Soft Computing* 24 (2014): 994-1012.

[126] Verbeke, Wouter, David Martens, and Bart Baesens. “Social network analysis for customer churn prediction.” *Applied Soft Computing* 14 (2014): 431-446.

[127] Xie, Haning, et al. “Predicting player disengagement and first purchase with event-frequency based data representation.” *2015 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 2015.

[128] Goldani, Mohammad Hadi, and Ali Goldani. “A review study on effective factors of developing the Finnish gaming industry and some suggestions for Iran’s game industry.” *2018 2nd National and 1st International Digital Games Research Conference: Trends, Technologies, and Applications (DGRC)*. IEEE, 2018.

[129] Mishra, Abinash, and U. Srinivasulu Reddy. “A novel approach for churn prediction using deep learning.” *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 2017.

[130] Umayaparvathi, V., and K. Iyakutti. “Automated feature selection and churn prediction using deep learning models.” *International Research Journal of Engineering and Technology (IRJET)* 4.3 (2017): 1846-1854.

[131] Burez, Jonathan, and Dirk Van den Poel. “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 (2007): 277-288.

[132] Burez, Jonathan, and Dirk Van den Poel. “Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department.” *Expert Systems with Applications* 35.1-2 (2008): 497-514.

[133] Madden, Gary, Scott J. Savage, and Grant Coble-Neal. “Subscriber churn in the Australian ISP market.” *Information economics and policy* 11.2 (1999): 195-207.

[134] Gerpott, Torsten J., Wolfgang Rams, and Andreas Schindler. “Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market.” *Telecommunications policy* 25.4 (2001): 249-269.

[135] Seo, DongBack, C. Ranganathan, and Yair Babad. “Two-level model of customer retention in the US mobile telecommunications service market.” *Telecommunications policy* 32.3-4 (2008): 182-196.

[136] Pendharkar, Parag C. “Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services.” *Expert Systems with Applications* 36.3 (2009): 6714-6720.

[137] Chu, Bong-Horng, Ming-Shian Tsai, and Cheng-Seen Ho. “Toward a hybrid data mining model for customer retention.” *Knowledge-Based Systems* 20.8 (2007): 703-718.

[138] Athanassopoulos, Antreas D. “Customer satisfaction cues to support market segmentation and explain switching behavior.” *Journal of business research* 47.3 (2000): 191-207.

[139] Kim, Hee-Su, and Choong-Han Yoon. “Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market.” *Telecommunications policy* 28.9-10 (2004): 751-765.